

Equilateral Non-Gaussianity and New Physics on the Horizon

Daniel Baumann and Daniel Green

School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540

Abstract

We examine the effective theory of single-field inflation in the limit where the scalar perturbations propagate with a small speed of sound. In this case the non-linearly realized time-translation symmetry of the Lagrangian implies large interactions, giving rise to primordial non-Gaussianities. When the non-Gaussianities are measurable, these interactions will become strongly coupled unless new physics appears close to the Hubble scale. Due to its proximity to the Hubble scale, the new physics is not necessarily decoupled from inflationary observables and can potentially affect the predictions of the model. To understand the types of corrections that may arise, we construct weakly-coupled completions of the theory and study their observational signatures.

Contents

1	Introduction	2
2	The Effective Theory of Inflation	5
2.1	Goldstone Description of Inflation	5
2.2	A Gauge Theory Analogy	7
2.3	Decoupling Limit	8
2.4	Energy Scales	8
2.5	A Hint of New Physics?	12
3	New Physics near Hubble	13
3.1	Preliminary Remarks	13
3.2	Weakly-Coupled UV-Completions	14
4	Observational Consequences	19
4.1	<i>Pessimistic View</i> : Intuitive Expectations	19
4.2	<i>Optimistic View</i> : Bispectra from Higher-Derivative Corrections	19
4.3	<i>Realistic View</i> : Observational Prospects	23
4.4	The $\omega_{\text{new}} \rightarrow 0$ Limit	24
5	Comments on Naturalness	26
6	Discussion	27
A	Strong Coupling in DBI Inflation	30
B	Dynamics of the π-σ Model	31
C	Corrections to Vanilla Sound Speed Models	33
C.1	Preliminaries	33
C.2	Two-Point Function	34
C.3	Three-Point Function	35
D	Small μ Limit of the π-σ Model	37
D.1	Canonical Quantization	37
D.2	Two-Point Function	38
D.3	Three-Point Function	39
E	Unitarity Bounds	40
E.1	Preliminaries	40
E.2	Unitarity and Small Sound Speed	40
E.3	The Extrinsic Curvature Model	41
E.4	The π - σ Model	42

1 Introduction

Effective field theory (EFT) [1, 2] is one of the most powerful organizing principles of all of theoretical physics. Its success is based on the basic observation that the physics at a particular scale of distance, time or energy doesn't depend sensitively on having detailed knowledge of the physics at widely different scales. The low-energy (or long-wavelength) degrees of freedom for the phenomena of interest can be isolated from the rest. EFT makes the procedure of eliminating unnecessary high-energy degrees of freedom precise, while systematically keeping track of their influence on the low-energy problem. The EFT approach has been applied successfully to virtually every area of theoretical physics, but its application to cosmology is rather recent, e.g. [3, 4, 5].

Given an effective theory that describes a set of experiments, one may ask when 'new physics' is expected to become important. Specifically, one would like to know what future experiments would require knowledge beyond the low-energy effective theory. In particle physics, future experiments typically involve colliders with sufficiently high center of mass energies E_{cm} to excite the new degrees of freedom. In that case, one would like to get a sense for the energy scale at which new particles are expected to be produced. A common procedure for identifying the scale of new physics is to determine the energy scale at which the effective theory becomes strongly coupled. As a concrete example, let us consider the Standard Model without the Higgs boson. The low-energy effective theory with generic W and Z couplings becomes strongly coupled—and WW scattering violates perturbative unitarity—when $E_{\text{cm}} > M_W/\sqrt{\alpha} \sim 1$ TeV [6]. We therefore expect the Higgsless effective theory to break down and some form of new physics to become important at (or below) that energy scale. Introducing a light Higgs particle, of course, eliminates the strong coupling at the TeV scale and yields an effective description that is, in principle, valid up to the Planck scale.

In this paper, we will explore an analogous situation in the context of inflationary cosmology [7]. Specifically, we will identify a regime in the effective theory of inflation [3] for which the leading interactions become strongly coupled not far above the Hubble scale H . As in the case of the Standard Model Higgs, this suggests that new physics becomes relevant at experimentally accessible energies.

Observations of the primordial density fluctuations, via their imprints in the cosmic microwave background (CMB), provide information about quantum-mechanical fluctuations of all fields that are lighter than the Hubble scale during inflation. Recently, Cheung et al. [3] (see also [4]) and Senatore and Zaldarriaga [5] developed effective theories which characterize these fluctuations and their interactions at horizon crossing, i.e. at energies near the Hubble scale. One of these effective theories is likely to provide a complete description of future experimental data. The theories are, in principle, valid descriptions not just at the Hubble scale but also at higher energies. In the following, we aim to understand how and when new physics is required to alter the effective theory of single-field inflation [3] when we extrapolate it to higher energies (extending our results to the effective theory of multi-field inflation [5] would be straightforward). Just like in the example of the Higgs, we will use strong coupling as a guide to new physics. Since large interactions in the effective theory give rise to large non-Gaussianities of the fluctuations [8], strong coupling bounds are most interesting for inflationary scenarios with measurable deviations from Gaussianity. We will find that observable equilateral non-Gaussianity implies a scale for the new physics that is

not far above the Hubble scale (see Figure 1). One may then hope that the new physics is not completely decoupled and can lead to subtle signatures in the data.¹

The effective theory of inflation [3] is based on the crucial insight that the inflationary perturbations are Goldstone bosons of spontaneously broken time-translation invariance of the quasi-de Sitter background. The curvature perturbations associated with these Goldstone modes lead to the temperature anisotropies observed in the CMB. Since Lorentz symmetry is broken by the time-dependent background, fluctuations may propagate with a velocity (‘speed of sound’) that is smaller than the speed of light, $c_s \leq 1$. Moreover, being Goldstone modes, the action for the perturbations is highly constrained by symmetry [1]. In particular, non-linearly realized time-translation symmetry relates a small value of c_s to large interactions and hence large equilateral non-Gaussianities, $f_{\text{NL}}^{\text{equil.}} \sim c_s^{-2}$ [3]. This observational signature is our prime motivation for a careful treatment of small speed of sound in the EFT of inflation.²

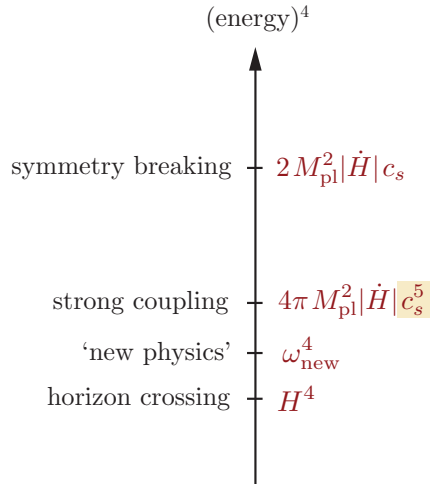


Figure 1: *Relevant energy scales in single-field inflation with small speed of sound.*

As we will show, the effective theory for single-field inflation with small c_s is characterized by a special hierarchy between the fundamental energy scales of the problem (see Figure 1): the symmetry breaking scale (Λ_b), the strong coupling scale (Λ_*) and the Hubble scale (H). Consistency of the effective theory requires H to be the lowest of these scales. The symmetry

¹At this point, it is worth remarking that the analogy between CMB and particle physics experiments is not perfect. For instance, one may rightly be concerned that future CMB experiments will not probe energy scales above Hubble directly. In that sense, it might seem that no cosmological experiment is sensitive to the new physics. While, in principle, this is a legitimate worry, in practice, the proximity of the scale of new physics to the Hubble scale allows us to learn about the new physics without producing the new degrees of freedom. This is similar to electroweak precision tests [9], which constrain Higgs physics without actually producing the Higgs particle.

²Understanding the backgrounds that give rise to a small speed of sound in models of inflation is a well-known theoretical challenge. Deriving a small speed of sound from a single scalar field requires that all orders in the derivative expansion of the background field ϕ are equally important [10, 11], yet be stable under radiative corrections [12]. This can only be achieved if the background respects a second symmetry—in addition to the shift symmetry of ϕ or the time-translation invariance of H —which protects the form of the higher-derivative interactions: In DBI inflation [12] the derivative expansion of the effective theory for the background is controlled by a higher-dimensional boost symmetry. In galileon inflation [13] the theory is protected by spacetime translation invariance. We will offer a few more comments about DBI inflation in Appendix A (for related thoughts see [14]).

breaking scale is the scale at which the background is integrated out, giving a theory of the fluctuations only. Therefore, if $\Lambda_\star \geq \Lambda_b$, the effective description changes before the theory becomes strongly coupled. This is indeed the case for slow-roll models of inflation. However, when $c_s \ll 1$, we find $\Lambda_\star^4 \sim c_s^4 \Lambda_b^4 \ll \Lambda_b^4$, so that there is a range of energies, $\Lambda_\star < \omega < \Lambda_b$, where the fluctuations appear to be strongly coupled and perturbative unitarity is lost. As in the case of particle physics, we expect new physics to become important at or below the scale where the effective theory becomes strongly coupled. In fact, if the theory is to remain weakly coupled at all energies, the new physics will become important at energies parametrically smaller than the strong coupling scale. The ratio $\omega_{\text{new}}/\Lambda_\star$ reflects an expansion parameter of the weakly-coupled completion of the effective theory.

Using the measured amplitude of the power spectrum of curvature fluctuations and the relation between c_s and the amplitude of non-Gaussianity $f_{\text{NL}}^{\text{equil.}}$, we find a strong upper bound on the scale of new physics ω_{new} (see §3):

$$\omega_{\text{new}} \ll \mathcal{O}(5) \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{-1/2} H. \quad (1.1)$$

Here, the use of “ \ll ” is meant to indicate that the theory is necessarily strongly coupled at the upper limit Λ_\star . We will be interested in models that are weakly coupled even above this energy scale. Therefore, the ratio $\omega_{\text{new}}/\Lambda_\star$ should be sufficiently small, so that the perturbative expansion is under control. Given that ω_{new} isn’t parametrically larger than the energy scale of inflation, we may hope that the effects of new physics aren’t completely decoupled at H and hence potentially observable.

The organization of the paper is as follows: In Section 2, we review the effective theory of inflation [3] and derive its relevant energy scales. We argue that models with small speed of sound require new physics to appear close to the Hubble scale. In Section 3, in the hopes of shedding light on the nature of this new physics, we explore weakly-coupled ultraviolet (UV)-completions³ of inflationary models with small speed of sound (see also [15, 16, 17, 18]). For each theory we determine the scale at which new physics becomes important. In Section 4, we compute possible observational signatures of our theories. As expected, for $\omega_{\text{new}} > H$, the leading bispectrum is predominantly equilateral and therefore indistinguishable from the strongly-coupled ‘pure c_s -theory’. However, we identify a subleading higher-derivative correction with a shape peaking both in the equilateral and the squashed momentum configurations. We discuss when this signature may be detectable. We also explain why this result motivates understanding the regime where the scale of new physics approaches the Hubble scale, $\omega_{\text{new}} \rightarrow H$, and the signal in the squashed configuration becomes an order-one contribution. In Section 5, we comment on some differences concerning the notion of naturalness in our weakly-coupled UV-completions relative to the strongly-coupled effective theories. We conclude with Section 6.

Five appendices contain important technical details: In Appendix A, we clarify the connection between the strong coupling scale in the effective theory of the fluctuations and properties of

³We will use the term “UV-completion” in a weaker sense than usual. We will call a theory “UV-complete”, if the effective theory of the fluctuations is weakly coupled up to the symmetry breaking scale, at which point the background becomes important. However, this does not mean that we have a theory for the background that would give rise to the effective theory of the fluctuations, nor have we embedded the theory in a UV-completion of gravity like string theory.

well-known UV-completions of the background, such as DBI inflation [12]. In Appendix B, we explain the dynamics of one of the UV-completions of Section 3 by examining the solutions to the equations of motion. In Appendix C, we provide details of the bispectrum calculation for our theories. In Appendix D, we present the bispectrum calculation in a novel model that arises when the scale of new physics is below the Hubble scale. Finally, in Appendix E, we explicitly compute the strong coupling scales for all models considered in this paper.

Throughout the text we employ the metric signature $(-, +, +, +)$ and use Greek letters μ, ν, \dots , to denote spacetime indices, while reserving Latin letters i, j, \dots , to label spatial indices. Spacetime indices are contracted with the metric $g_{\mu\nu}$, while spatial indices are contracted with the Kronecker delta δ_{ij} . We choose natural units with $c = \hbar = 1$ and define the reduced Planck mass as $M_{\text{pl}} = (8\pi G)^{-1/2}$. We use the letter π to denote both $3.141\dots$ and the Goldstone boson of broken time-translations. Which is meant should be clear from the context.

2 The Effective Theory of Inflation

The effective theory of inflation [3] provides a powerful and unified way of characterizing single-field models of inflation. It crucially exploits the fact that inflation spontaneously breaks time-translation symmetry. In this section we review the effective action of the Goldstone boson associated with the spontaneous symmetry breaking. We will explain that the theory is highly constrained by the non-linearly realized symmetries of the quasi-de Sitter background. For instance, by symmetry, a small speed of sound also implies large interactions. We will derive the strong coupling scale associated with these interactions and show that it isn't far above the Hubble scale if non-Gaussianity is to be observable.

2.1 Goldstone Description of Inflation

The construction of the effective theory of inflation [3] begins in ‘unitary gauge’, in which there are no matter fluctuations, but only metric fluctuations. The most general effective action is then constructed by writing down all operators that are functions of the metric fluctuations and invariant under time-dependent spatial diffeomorphisms. The two most important objects appearing in this construction are the metric perturbation $\delta g^{00} = g^{00} + 1$ and the extrinsic curvature perturbation $\delta K_{\mu\nu} = K_{\mu\nu} - a^2 H h_{\mu\nu}$, where $h_{\mu\nu}$ is the induced metric on the spatial slices. We use these geometrical quantities to write down the most general action with unbroken spatial diffeomorphisms [3]

$$\begin{aligned}
S = \int d^4x \sqrt{-g} & \left[\frac{1}{2} M_{\text{pl}}^2 R + M_{\text{pl}}^2 \dot{H} g^{00} - M_{\text{pl}}^2 (3H^2 + \dot{H}) \right. \\
& + \frac{1}{2!} M_2^4(t) (\delta g^{00})^2 + \frac{1}{3!} M_3^4(t) (\delta g^{00})^3 + \dots \\
& \left. - \frac{1}{2} \bar{M}_1^3(t) \delta g^{00} \delta K_\mu^\mu - \frac{1}{2} \bar{M}_2^2(t) (\delta K_\mu^\mu)^2 - \frac{1}{2} \bar{M}_3^2(t) \delta K^{\mu\nu} \delta K_{\mu\nu} + \dots \right]. \quad (2.1)
\end{aligned}$$

Since time diffeomorphisms are broken by the time-dependence of the background, the ‘couplings’ $H(t)$, $M_n(t)$ and $\bar{M}_n(t)$ are functions of time. However, the time-translation invariance is only weakly broken during inflation, so the time-dependence of these functions is typically small—e.g. $|\dot{H}| \ll H^2$. As usual in effective field theories, the action is organized as a low-energy

expansion of the fields and their derivatives: g^{00} is a scalar with zero derivatives acting on it, while $K_{\mu\nu}$ is a one-derivative object. In many situations the terms involving g^{00} therefore dominate the dynamics.

Let us show how the effective action (2.1) unifies a large class of single-field models of inflation:⁴

- The first line in (2.1) captures all single-field slow-roll models of inflation

$$\mathcal{L}_{\text{sr}} = -\frac{1}{2}(\partial_\mu\phi)^2 - V(\phi) \rightarrow \frac{1}{2}\dot{\phi}^2 g^{00} - V(\phi) \Leftrightarrow \frac{1}{2}\dot{\phi}^2 = M_{\text{pl}}^2 |\dot{H}|. \quad (2.2)$$

- The second line parameterizes models with non-trivial kinetic terms

$$\mathcal{L}_{\text{p(x)}} = P(X, \phi) \rightarrow P(\dot{\phi}^2 g^{00}, \bar{\phi}) \Leftrightarrow M_n^4 = \dot{\phi}^{2n} \frac{\partial^n P}{\partial \bar{X}^n}. \quad (2.3)$$

The operators proportional to M_n^4 start at order n in the fluctuations. The coefficient M_2 induces a sound speed in the quadratic action

$$c_s^{-2} \equiv 1 - \frac{2M_2^4}{M_{\text{pl}}^2 \dot{H}}. \quad (2.4)$$

- The last line in (2.1) characterizes terms with higher derivatives that cannot be eliminated by partial integrations, such as $(\Box\phi)^2$. Typically, these terms are suppressed by extra powers of the cutoff of the theory. However, they can become important in cases like ghost inflation [22] and its generalizations [23], where the leading terms vanish because $M_{\text{pl}}^2 \dot{H} \rightarrow 0$.

The power of the approach of [3] is that it is a completely general description for any background $H(t)$ that spontaneously breaks time-translation invariance. In particular, it is independent of microscopic details of the theory that gives rise to the de Sitter background. However, as written, the dynamics of the theory are not at all clear. To make the dynamics more transparent, we introduce the Goldstone boson π associated with the spontaneous breaking of time-translation invariance. Moreover, via the Stückelberg trick, π restores the full gauge-invariance of the theory. Specifically, by definition, the Goldstone transforms as $\pi \rightarrow \pi - \xi(x, t)$ under the time reparameterization $t \rightarrow t + \xi(x, t)$, such that $t + \pi$ is invariant. With the replacements $t \rightarrow t + \pi$ and $g^{00} \rightarrow \partial_\mu(t + \pi)\partial_\nu(t + \pi)g^{\mu\nu}$ the action (2.1) becomes fully gauge-invariant. For example, the slow-roll and $P(X)$ parts of (2.1) are given by

$$\mathcal{L}_{\text{sr}} = M_{\text{pl}}^2 \dot{H}(t + \pi) \left[\partial_\mu(t + \pi)\partial_\nu(t + \pi)g^{\mu\nu} \right] - M_{\text{pl}}^2 (3H^2 + \dot{H})(t + \pi), \quad (2.5)$$

$$\mathcal{L}_{\text{p(x)}} = \frac{1}{2}M_2^4(t + \pi) \left[\partial_\mu(t + \pi)\partial_\nu(t + \pi)g^{\mu\nu} + 1 \right]^2 + \dots. \quad (2.6)$$

The quadratic term in (2.6) modifies the kinetic term for the Goldstone boson, but not the gradient-squared term. It therefore leads to the sound speed cited in (2.4). From (2.6) we observe that the non-linearly realized symmetry relates a small c_s (large M_2) to large interactions and

⁴The formalism does not capture inflationary models with significant dissipative effects, such as [19, 20]. An extension of the effective theory of inflation that incorporates dissipation will appear in [21].

hence observable non-Gaussianity. This limit will be of particular interest to the considerations in this paper.

Inflationary observables are often expressed in terms of the conserved curvature perturbation on comoving slices⁵, ζ . Performing a temporal gauge transformation, we relate the Goldstone boson to the comoving curvature perturbation, $\zeta = -H\pi$. Hence, up to corrections suppressed by $\frac{\dot{H}}{H^2}$, the correlation functions of π are proportional to the correlation functions of ζ .

2.2 A Gauge Theory Analogy

The procedure of reintroducing the Goldstone boson is common in the description of massive vector bosons. Since we will make frequent use of this analogy, we digress briefly to review the gauge theory example. Consider a non-Abelian gauge theory with Lagrangian

$$\mathcal{L} = -\frac{1}{4} \text{Tr} F_{\mu\nu}^2 - \frac{1}{2} m^2 \text{Tr} A_\mu^2 . \quad (2.7)$$

Under a gauge transformation with

$$A_\mu \rightarrow U A_\mu U^\dagger + \frac{i}{g} U \partial_\mu U^\dagger \equiv \frac{i}{g} U D_\mu U^\dagger , \quad (2.8)$$

the action becomes

$$\mathcal{L} = -\frac{1}{4} \text{Tr} F_{\mu\nu}^2 - \frac{1}{2} \frac{m^2}{g^2} \text{Tr} D_\mu U^\dagger D^\mu U . \quad (2.9)$$

Gauge invariance is restored by the Stückelberg trick, i.e. defining $U = e^{i\pi^a T^a}$, where T^a is a generator of the group. Choosing unitary gauge, $\pi^a \equiv 0$, reproduces the action in (2.7). Including π , the Lagrangian becomes gauge-invariant and can be expanded as

$$\mathcal{L} = -\frac{1}{4} \text{Tr} F_{\mu\nu}^2 - \frac{1}{2} m^2 \text{Tr} A_\mu^2 + \frac{1}{2} \frac{m^2}{g^2} (\partial_\mu \pi^a)^2 + i \frac{m^2}{g} \text{Tr} \partial_\mu \pi^a T^a A^\mu + c.c. + \dots . \quad (2.10)$$

One of the main advantages of including the Goldstone bosons is that it makes the high-energy behavior of the theory manifest. Specifically, it tells us that at high energies, the scattering of the longitudinal mode of the gauge field is well-described by the scattering of the Goldstone bosons. This is most easily seen by taking the decoupling limit $m \rightarrow 0$ and $g \rightarrow 0$, while keeping $m/g \equiv f_\pi$ fixed. In this limit, the Goldstone bosons decouple from A_μ and we are left with

$$\mathcal{L} = -\frac{1}{2} f_\pi^2 \text{Tr} \partial_\mu U^\dagger \partial^\mu U . \quad (2.11)$$

Restoring finite m and g , we should expect corrections to the results from pure Goldstone boson scattering that are perturbative in m/E and g^2 , where E is the energy of the vector boson.

⁵When it was first introduced this quantity was called \mathcal{R} , in order to distinguish it from the curvature perturbation on uniform density slices, ζ . Here, we follow [24] and (mis)use ζ for the comoving curvature perturbation.

2.3 Decoupling Limit

In inflation, we would similarly like to understand the circumstances under which π alone controls the behavior of correlation functions. Specifically, we want to quantify the error that is made when correlation functions are computed from the *decoupled* π -lagrangian \mathcal{L}_π , i.e. the part of the Lagrangian that doesn't include the mixing with metric fluctuations. We can decouple the Goldstone boson π from gravitational fluctuations by taking the limit $M_{\text{pl}} \rightarrow \infty$ and $\dot{H} \rightarrow 0$, with $M_{\text{pl}}^2 \dot{H}$ fixed. This limit is equivalent to the gauge theory example via the following identifications: $M_{\text{pl}}^{-1} \leftrightarrow g$, $\dot{H} \leftrightarrow m^2$ and $H \leftrightarrow E$. Therefore, if we compute correlation functions using the decoupled π -lagrangian, our answers should be accurate up to fractional corrections of order $\frac{H^2}{M_{\text{pl}}^2}$ and $\frac{\dot{H}}{H^2} \equiv -\epsilon$.

We can also argue this by considering the dynamics of the comoving curvature perturbation $\zeta = -H\pi$. During single-field inflation, the curvature perturbation freezes on superhorizon scales, $\dot{\zeta} \rightarrow 0$ [25]. In other words, ζ is massless outside of the horizon. From $\dot{\zeta} = -H\dot{\pi} - \dot{H}\pi$ this implies that π cannot be precisely massless. However, since π is indeed massless in the decoupling limit, this mass for π must be coming from the mixing with gravity. To compensate for the time-dependence of the Hubble rate in the relation $\zeta = -H(t)\pi$, the terms associated with the mixing with gravity such as $\delta N \dot{\pi}$ and $N^i \partial_i \pi$, where N and N_i are the standard ADM variables [26], must be proportional to \dot{H} .⁶ Note that the time-dependence of π outside the horizon is independent of the dispersion relation. Indeed, solving the constraint equations implied by (2.5) and (2.6), we find [27]

$$\delta N = \epsilon H \pi \quad \text{and} \quad \partial^i N_i = -\frac{\epsilon H \dot{\pi}}{c_s^2} . \quad (2.12)$$

Plugging this back into the quadratic action gives

$$\mathcal{L} = -\frac{M_{\text{pl}}^2 \dot{H}}{c_s^2} \left(\dot{\pi}^2 - \frac{c_s^2}{a^2} (\partial_i \pi)^2 + 3\epsilon H^2 \pi^2 \right) . \quad (2.13)$$

This explicitly confirms that the mass for π arising from the mixing with gravity is $m_\pi^2 = 3\epsilon H^2 = -3\dot{H}$. In the remainder we will restrict to the decoupling limit,

$$g^{00} \rightarrow -(1 + \dot{\pi})^2 + \frac{(\partial_i \pi)^2}{a^2} , \quad (2.14)$$

in which the π -lagrangian becomes

$$\mathcal{L} \rightarrow -\frac{M_{\text{pl}}^2 \dot{H}}{c_s^2} \left[\dot{\pi}^2 - \frac{c_s^2}{a^2} (\partial_i \pi)^2 - (1 - c_s^2) \left(\dot{\pi}^3 - \dot{\pi} \frac{(\partial_i \pi)^2}{a^2} \right) \right] + \dots . \quad (2.15)$$

2.4 Energy Scales

In order to understand the dynamics of a model, it is important to identify the energy scales at which different phenomena become important. Three energy scales are particularly relevant in the effective theory of inflation:

⁶At least the terms that give π a time variation outside the horizon must be proportional to \dot{H} . In principle, mixing with gravity could also change derivative operators where the size of the effect cannot be estimated in this way.

- the *symmetry breaking scale*, Λ_b , is the energy scale at which time translations are spontaneously broken and a description in terms of a Goldstone boson first becomes applicable;
- the *strong coupling scale*, Λ_* , defines the energy scale at which the effective description breaks down and perturbative unitarity is lost;
- the *Hubble scale*, H , is the energy scale associated with the cosmological experiment.

In slow-roll inflation, these three energy scales can easily be identified. Time-translation invariance is broken by the background $\bar{\phi}(t)$ at the scale $\Lambda_b^2 = \dot{\bar{\phi}}$. At energy scales above Λ_b , the symmetry is restored and we should not integrate out the background. Because the theory is effectively Gaussian, the self-interactions of ϕ are weak up to very high energies. The theory only becomes strongly coupled at the Planck scale, so the UV-cutoff is M_{pl} . Inflationary observables freeze out at horizon-crossing, or when their frequencies become equal to the expansion rate, $\omega \sim H$. Inflation therefore directly probes energies of order Hubble, i.e. the energy scale of the experiment is H . We will now define these energy scales rigorously in the effective theory, so that they can be identified in models other than slow-roll inflation. Readers who don't want to follow the details of the derivations may jump directly to §2.5 where we summarize the results and discuss their implications.

2.4.1 Symmetry Breaking

Although the inflationary background spontaneously breaks a gauge symmetry, in the decoupling limit the gauge symmetry becomes a global symmetry. As long as the decoupled π -lagrangian is a reliable description, the language of spontaneously broken global symmetries will therefore be useful. In this section we will formulate the effective theory of inflation in a way that makes this analogy manifest.

Let us first review the more familiar case of a spontaneously broken internal symmetry, i.e. the decoupled Goldstone boson in a conventional gauge theory. By definition, any theory with a continuous global symmetry has a conserved Noether current J^μ even if the symmetry is spontaneously broken. There may also be an associated conserved charge

$$Q = \int d^3x J^0(x) . \quad (2.16)$$

The existence of a well-defined Q requires $J^0(x)$ to vanish at least as x^{-3} in the limit $x \rightarrow \infty$. In momentum space, this means that $J^0(p)$ scales at most like p^{-1} for $p \rightarrow 0$. When the symmetry is spontaneously broken there is no conserved charge at infinity. While the current always exists and satisfies $\partial_\mu J^\mu = 0$, the charge itself is not well-defined. Therefore, when the global symmetry is spontaneously broken, $J^0(x)$ will have contributions that do not fall off sufficiently rapidly as $x \rightarrow \infty$. Nevertheless, even when the charge at infinity diverges, commutators of local fields with the charges are still well-defined.

The current associated with the non-linear sigma model in (2.11) is

$$J^\mu = -f_\pi \partial^\mu \pi_c + \cdots , \quad \text{where} \quad f_\pi \equiv \frac{m}{g} . \quad (2.17)$$

Here we have defined the canonically-normalized Goldstone field $\pi_c \equiv f_\pi \pi$. The normalization of the current (2.17) is consistent with $[Q, \pi] = 1 + \dots$ (or $[Q, \pi_c] = f_\pi + \dots$) and the commutator of the canonically-normalized field, $[\tilde{\pi}_c(\mathbf{x}), \pi_c(\mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y})$. As $x \rightarrow \infty$, we expect $\partial_0 \pi \sim x^{-2}$ by dimensional analysis and the charge therefore does not exist. This is a direct consequence of the fact that the π field shifts under the symmetry.

After these preliminary remarks, we will now use the current of the non-linear sigma model (2.17) to determine the scale at which the symmetry is broken. We find the following two-point function⁷

$$\int d^4x e^{ipx} \langle 0 | T \{ J^\mu(x) J^\nu(0) \} | 0 \rangle = i(p^\mu p^\nu - \eta^{\mu\nu} p^2) \Pi(p^2) , \quad (2.18)$$

where

$$\Pi(p^2) \equiv \frac{f_\pi^2}{p^2} + \mathcal{O}(1) . \quad (2.19)$$

The first term in (2.19) implies $J^0 \sim p^{-2}$ and therefore the charge at infinity does not exist. As a result, the symmetry is spontaneously broken at low energies. The higher-order terms in (2.19) are terms that are consistent with the symmetry being unbroken. Therefore, when $p^2 \gg f^2$, the higher-order terms dominate and the symmetry appears to be unbroken.

Returning to inflation, we would like to determine from (2.15) the scale at which time-translation invariance is spontaneously broken. To exploit the analogy with gauge theory, we reintroduce (fake) Lorentz invariance of the action by rescaling the spatial coordinates:

$$\tilde{\mathcal{L}} = -\frac{1}{2}(\tilde{\partial}_\mu \tilde{\pi}_c)^2 + \dots , \quad (2.20)$$

where $x \rightarrow \tilde{x} \equiv c_s^{-1}x$ and $\mathcal{L} \rightarrow \tilde{\mathcal{L}} = c_s^3 \mathcal{L}$. Here we have defined the canonically-normalized field

$$\tilde{\pi}_c^2 \equiv (2M_{\text{pl}}^2 |\dot{H}| c_s) \pi^2 . \quad (2.21)$$

The Noether current associated with (2.20) is

$$\tilde{J}^\mu \equiv \tilde{T}^{\mu 0} = -f^2 \tilde{\partial}^\mu \tilde{\pi}_c + \dots , \quad \text{where} \quad f^2 \equiv (2M_{\text{pl}}^2 |\dot{H}| c_s)^{1/2} . \quad (2.22)$$

The normalization of the current is consistent with $[Q, \pi] = -1$, where $Q \equiv \int d^3\tilde{x} \tilde{T}^{00}$. Since we rescaled the spatial momenta such that everything is an energy scale, we can read off the symmetry breaking scale from our previous discussion of global currents. We conclude that the symmetry is spontaneously broken at

$$\Lambda_b^4 \equiv 2M_{\text{pl}}^2 |\dot{H}| c_s . \quad (2.23)$$

One can arrive at the same result without rescaling the spatial coordinates x by being careful to define an energy scale. Recall that $T^{00} = 2M_{\text{pl}}^2 \dot{H} c_s^{-2} \dot{\pi}^2 + \dots$ is an energy density, i.e. an energy over volume. Because $\dot{\pi}$ is dimensionless, $2M_{\text{pl}}^2 \dot{H} c_s^{-2}$ must have units of $[\omega][k]^3$. To discuss energies, we use the dispersion relation, $\omega = c_s k$, to find $\Lambda_b^4 \equiv c_s^3 (2M_{\text{pl}}^2 |\dot{H}| c_s^{-2}) = 2M_{\text{pl}}^2 |\dot{H}| c_s$ as

⁷We have chosen contact terms such that J^μ is conserved when $x = 0$. This choice is necessary if one is weakly gauging the symmetry.

before. Using the dispersion relation to define an energy in this way will be useful in cases where the rescaling of x is ineffective (cf. §3.2.2).

Although the current gives a natural definition of the symmetry breaking scale, it is nice to check that it agrees with our intuition. First of all, in the case of slow-roll inflation (i.e. for $c_s = 1$), the symmetry breaking scale is given by $2M_{\text{pl}}^2|\dot{H}| = \dot{\phi}^2$. This matches the intuition that the time variation of the background is breaking the symmetry. Moreover, one may rewrite the (dimensionless) power spectrum of curvature fluctuations (cf. §4) in terms of the symmetry breaking scale

$$\Delta_\zeta \equiv k^3 P_\zeta(k) = \frac{1}{4} \frac{H^2}{M_{\text{pl}}^2 \epsilon c_s} = \frac{1}{2} \left(\frac{H}{\Lambda_b} \right)^4. \quad (2.24)$$

Hence, when $H \sim \Lambda_b$, the size of quantum fluctuations is of the same order as the symmetry breaking scale. This is the regime of *eternal inflation*, which is again consistent with the interpretation of $\Lambda_b^4 = \dot{\phi}^2$ for slow-roll.

2.4.2 Strong Coupling

The regime of validity of an effective theory is not always obvious. Given a microscopic definition of the theory (i.e. a UV-completion), the regime of validity is determined by the scales at which additional modes were integrated out. Given only the effective description, these energy scales may not be transparent in the Lagrangian. A fairly reliable method to identify the cutoff of the effective theory is to determine the energy scale at which the theory becomes strongly coupled. This is the approach that we will follow. Many of the results of this section are derived in detail in Appendix E.

Let us again use the theory of massive gauge bosons as an example. Given the action in unitary gauge (2.7), it is not a priori clear where the effective description breaks down. By carefully studying the behavior of scattering amplitudes, one finds that the scattering of the longitudinal modes of the gauge fields becomes strongly coupled at the scale $4\pi\Lambda_\star^2 = 4\pi m^2/g^2$. This becomes more transparent after we introduce the Goldstone bosons in (2.10). The action is an expansion in π_c/f_π which contains irrelevant operators of arbitrarily large dimensions. By dimensional analysis, the effective coupling is ω/f_π which makes the strong coupling scale at $\omega^2 = 4\pi f_\pi^2$ manifest.

Returning to the effective theory of inflation, we can similarly determine the strong coupling scale from the action for the Goldstone boson [3]. For a general speed of sound, we first rescale the spatial coordinates as we did in (2.20). The non-linear realization of the time-translation symmetry enforces relations between the quadratic, cubic and quartic actions. Keeping only the leading interactions, we find

$$\tilde{\mathcal{L}} = -\frac{1}{2}(\tilde{\partial}_\mu \tilde{\pi}_c)^2 + \frac{1}{2} \frac{(1 - c_s^2)}{\Lambda_\star^2} \dot{\tilde{\pi}}_c \frac{(\tilde{\partial}_i \tilde{\pi}_c)^2}{a^2} + \frac{1}{8} \frac{1}{\Lambda_\star^4} \frac{(\tilde{\partial}_i \tilde{\pi}_c)^4}{a^4}, \quad (2.25)$$

where

$$\Lambda_\star^4 \equiv 2M_{\text{pl}}^2 |\dot{H}| c_s^5 (1 - c_s^2)^{-1}. \quad (2.26)$$

As in the gauge theory example, the effective coupling is given by ω/Λ_\star . We expect that strong coupling arises at some order-one value of this coupling. It is useful to define the strong coupling

scale by the breakdown of perturbative unitarity of the Goldstone boson scattering. We calculate this scale in Appendix E and find that the theory is strongly coupled when $\omega^4 = 2\pi\Lambda_\star^4$. Note the large suppression of Λ_\star^4 by factors of $c_s \ll 1$. Without rescaling the coordinates, the factors of c_s in the strong coupling scale are less obvious. However, the powers of c_s will always agree because they convert momentum scales into energy scales. As a result, the powers of c_s are uniquely determined when we write the strong coupling scale as an energy scale. We will explain this in more detail in §3.

The interactions which become strongly coupled are the same that give rise to measurable non-Gaussianity. As a result, we should be able to interpret the strong coupling scale in terms of the size of the non-Gaussianity. A simple estimate for the amplitude of the non-Gaussianity is

$$f_{\text{NL}} \zeta \equiv \left. \frac{\mathcal{L}_3}{\mathcal{L}_2} \right|_{\omega=H} \sim \frac{M_{\text{pl}}^2 \dot{H} c_s^{-2} (1 - c_s^2) \dot{\pi} \frac{(\partial_i \pi)^2}{a^2}}{M_{\text{pl}}^2 \dot{H} \dot{\pi}^2} = c_s^{-2} (1 - c_s^2) \zeta \sim \left(\frac{\Lambda_{\text{b}}}{\Lambda_\star} \right)^2 \zeta. \quad (2.27)$$

Using the power spectrum (2.24) as an estimate for the size of $\zeta \sim \Delta_\zeta^{1/2}$, we find

$$\frac{\mathcal{L}_3}{\mathcal{L}_2} \sim \left(\frac{H}{\Lambda_\star} \right)^2. \quad (2.28)$$

We see that $\mathcal{L}_3 \sim \mathcal{L}_2$, or $f_{\text{NL}} \sim \zeta^{-1} \sim 10^4$, when $H \sim \Lambda_\star$. This indicates a breakdown of the perturbative description as Λ_\star approaches H .

2.5 A Hint of New Physics?

Summary. In the previous sections we derived two important energy scales in the effective theory of inflation: the symmetry breaking scale, $\Lambda_{\text{b}}^4 = 2M_{\text{pl}}^2 |\dot{H}| c_s$, and the strong coupling scale, $\Lambda_\star^4 = 2M_{\text{pl}}^2 |\dot{H}| c_s^5 (1 - c_s^2)^{-1}$. In slow-roll inflation, $c_s \rightarrow 1$, the strong coupling scale is much larger than the symmetry breaking scale. However, in models with small speed of sound, $c_s \ll 1$, this hierarchy of scales is reversed,

$$\frac{\Lambda_{\text{b}}^4}{\Lambda_\star^4} = (1 - c_s^2) c_s^{-4} \simeq 16 (f_{\text{NL}}^{\text{equil.}})^2, \quad (2.29)$$

where we used (2.27) to relate c_s to f_{NL} , or more precisely $f_{\text{NL}}^{\text{equil.}} \approx -(4c_s^2)^{-1}$ (see §4). Therefore, any measurable non-Gaussianity ($f_{\text{NL}}^{\text{equil.}} \gtrsim 10$) requires the strong coupling scale to appear parametrically below the scale at which the background is integrated out.

Implications. The inherently strongly-coupled nature of the above theories was a result of restricting the particle content of the model. However, just like in particle physics, one should take this as an indication that new degrees of freedom may become important at energies below the scale of strong coupling.⁸ Therefore, there is an energy scale ω_{new} at which ‘new physics’ becomes important. If we wish to maintain both weak coupling and the effective small c_s -description at Hubble, we require $H^2 < \omega_{\text{new}}^2 \ll \sqrt{2\pi} \Lambda_\star^2$. Given our previous results, we find

$$\frac{H^4}{\Lambda_\star^4} = 32 \Delta_\zeta (f_{\text{NL}}^{\text{equil.}})^2. \quad (2.30)$$

⁸The ‘new physics’ could also take the form of a change in the physical description of the existing degrees of freedom. In fact, this is the case in our UV-completions in §3.

where $\frac{\Delta_\zeta}{2\pi^2} = 2.4 \times 10^{-9}$ and $|f_{\text{NL}}^{\text{equil.}}| \lesssim 300$ [28]. This implies that the new physics must enter not far above the Hubble scale:

$$H^2 < \omega_{\text{new}}^2 \ll \sqrt{2\pi} \Lambda_\star^2 \approx \mathcal{O}(20) \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{-1} H^2. \quad (2.31)$$

This range of energies is sufficiently small that the new physics is not obviously decoupled at the Hubble scale. The use of “ \ll ” in (2.31) is a reminder that our loop expansion is being controlled by the ratio $\omega^2/\sqrt{2\pi}\Lambda_\star^2$. When $\omega_{\text{new}}^2 \rightarrow \sqrt{2\pi}\Lambda_\star^2$ the theory becomes strongly coupled. However, quantum corrections of any observable become increasingly important as one approaches this limit. Therefore, a useful perturbative description requires that we expand in small $\omega^2/\sqrt{2\pi}\Lambda_\star^2$, to ensure that our description is not dominated by strong dynamics. This ratio reflects the strength of a coupling in any UV-completion of the effective theory.

3 New Physics near Hubble

In this section we will construct theories which at low energies look like small speed of sound models, $\mathcal{L}_{\text{sr}} + \frac{1}{2}M_2^4(\delta g^{00})^2$, but experience a change in their physical description at $\omega_{\text{new}}^2 \ll \sqrt{2\pi}\Lambda_\star^2$, such that they remain weakly coupled up to the symmetry breaking scale Λ_{b} .

3.1 Preliminary Remarks

The change in the physical description will not necessarily require new propagating degrees of freedom to enter at ω_{new} . Instead, it may simply be the case that the scaling behavior of the field changes. Indeed, we will find weakly-coupled theories, in which the dispersion relation changes from $\omega = c_s k$ to $\omega = k^2/\rho$ at $\omega_{\text{new}} = \rho c_s^2$, where ρ is some energy scale. The reason that this change in the dispersion relation modifies our previous result can be seen as follows: in a relativistic theory, an operator in the action of the form $\frac{1}{\Lambda^n} \int dt d^3x \mathcal{O}_{4+n}$ implies that the strong coupling scale is Λ if \mathcal{O}_{4+n} is constructed from canonically-normalized fields. However, in a non-relativistic theory, the coupling Λ first needs to be written as an ‘energy scale’. In order to do this, we have to use the dispersion relation. For example, in the action

$$S = \int dt d^3x \frac{a^3}{2} \left[\dot{\pi}_c^2 - \frac{c_s^2}{a^2} (\partial_i \pi_c)^2 + \frac{1}{\Lambda^2} \dot{\pi}_c \frac{(\partial_i \pi_c)^2}{a^2} \right], \quad (3.1)$$

the field π_c has units $[k]^{3/2}[\omega]^{-1/2}$ and therefore Λ^4 has units $[k]^7[\omega]^{-3}$. The linear dispersion relation $\omega = c_s k$ implies that the theory becomes strongly coupled at the energy scale: $\Lambda_\star^4 = \Lambda^4 c_s^7$. For small c_s , the strong coupling scale Λ_\star is therefore highly suppressed relative to the parameter Λ . Of course, this method for determining the strong coupling scale is equivalent to the rescaling procedure $x = c_s \tilde{x}$ of the previous section.

From this argument it becomes clear that by changing the dispersion relation, one can change the energy scale at which the theory becomes strongly coupled, even without changing the coefficient of the operator itself. Specifically, if the dispersion relation becomes non-linear before the would-be strong coupling scale is reached, $\omega \rightarrow k^2/\rho$, then the relation between the new strong coupling scale Λ_\star and the coupling Λ is $\Lambda_\star^4 = \Lambda^4(\Lambda/\rho)^{28}$. For $\rho^4 \ll \Lambda^4 c_s^{-1}$ this implies a larger strong coupling scale than in the case with linear dispersion relation. We will now present two

explicit examples of weakly-coupled UV-completions that implement this change in the dispersion relation.

3.2 Weakly-Coupled UV-Completions

3.2.1 The π - σ Model

A natural way to realize the new physics at ω_{new} is by coupling π to an additional degree of freedom⁹ σ . In unitary gauge we therefore consider the action

$$\mathcal{L} = \mathcal{L}_{\text{sr}} - \frac{1}{2}(\partial_\mu \sigma)^2 - m^{4-n}(g^{00} + 1)\mathcal{O}_n(\sigma) - V(\sigma) , \quad (3.2)$$

where $\mathcal{O}_n(\sigma)$ is an operator of dimension n involving only the field σ . The potential $V(\sigma)$ parameterizes the self-interactions of σ . Additional couplings $(g^{00} + 1)^m \mathcal{O}_n(\sigma)$ won't affect the quadratic action for the Goldstone π , but may be added when considering its interactions. We could also add derivative interactions of σ [5].

Our goal is to generate $c_s \ll 1$ by integrating out σ . For $n > 1$, one expects the kinetic term for π to be modified only by loops of σ . Therefore, let us choose $\mathcal{O}(\sigma) = \sigma$ and $V(\sigma) = \frac{1}{2}\mu^2\sigma^2$. Introducing the Goldstone boson and taking the decoupling limit, the action becomes

$$\mathcal{L} = -M_{\text{pl}}^2 |\dot{H}| (\partial_\mu \pi)^2 - \frac{1}{2}(\partial_\mu \sigma)^2 - J(\pi)\sigma - \frac{1}{2}\mu^2\sigma , \quad (3.3)$$

where

$$J(\pi) \equiv m^3(g^{00} + 1) \rightarrow m^3[-2\dot{\pi} + (\partial_\mu \pi)^2] . \quad (3.4)$$

Small sound speed. For certain ranges of the parameters m and μ (to be determined momentarily) we can integrate out σ —i.e. $e^{iS_{\text{eff}}(\pi)} = \int \mathcal{D}\sigma e^{iS[\pi, \sigma]}$ —to get an effective action for the field π ,

$$\mathcal{L}_{\text{eff}} = -M_{\text{pl}}^2 |\dot{H}| (\partial_\mu \pi)^2 + \frac{1}{2} J(\pi) (\mu^2 - \square)^{-1} J(\pi) , \quad (3.5)$$

where $\square \equiv g^{\mu\nu} \nabla_\mu \nabla_\nu \rightarrow -\partial_t^2 + \frac{1}{a^2} \partial_i^2 - 3H\partial_t$. Clearly, for some range of energies this is a non-local action. However, at sufficiently low energies a derivative expansion provides a useful description

$$\mathcal{L}_{\text{eff}} = -M_{\text{pl}}^2 |\dot{H}| (\partial_\mu \pi)^2 + \frac{2m^6}{\mu^2} \dot{\pi}^2 + \mathcal{O}\left(\frac{\square}{\mu^2}\right) + \dots . \quad (3.6)$$

This theory has a speed of sound given by

$$c_s^{-2} = 1 + \frac{\rho^2}{\mu^2} , \quad \text{where} \quad \rho^2 \equiv \frac{2m^6}{M_{\text{pl}}^2 |\dot{H}|} . \quad (3.7)$$

The expansion in derivatives is under control as long as $|\omega^2 - k^2| < \mu^2$. Since the dispersion relation is $k^2 = \omega^2 c_s^{-2}$, the expansion is reliable when

$$\omega^2 < \mu^2 c_s^2 \equiv \omega_{\text{new}}^2 . \quad (3.8)$$

⁹These theories are examples of the multi-field effective theory of Senatore and Zaldarriaga [5]. However, we will couple π to fields σ that are heavier than the Hubble scale, so we don't have to impose symmetries on the σ fields to make them naturally light. Different types of couplings will therefore be relevant in our case. Our theories will be equivalent to the effective theory for the fluctuations in the models of [15, 16, 17].

The speed of sound can also be derived directly from the equations of motion (see Appendix B). Although we have spoken of integrating out σ , we will now explain that *no* new particle appears at ω_{new} .¹⁰ Instead, ω_{new} simply marks the energy scale at which the dispersion relation changes from linear, $\omega = c_s k$, to non-linear, $\omega = k^2/\rho$.

Relation to previous works. A number of previous works have recognized the possibility of using multi-field dynamics to generate a small sound speed at low energies: e.g. in the *gelaton* model of Tolley and Wyman [15] a heavy field is strongly coupled to the kinetic energy of the inflation, leading to a curved inflaton trajectory [15, 16, 17]. Integrating out the heavy gelaton field leads to an effective single-field theory with small speed of sound. Here, we have shown how the effective theory for the fluctuations of these models arises from a systematic application of the approach of Senatore and Zaldarriaga [5]. Next, we clarify the dynamics of these theories.

Dynamics. To discuss the dynamics of the theory at all energies between H and $2M_{\text{pl}}^2|\dot{H}|$, we consider the quadratic part of the two-field action

$$\mathcal{L}_2 = -\frac{1}{2}(\partial_\mu \pi_c)^2 - \frac{1}{2}(\partial_\mu \sigma)^2 + \rho \dot{\pi}_c \sigma - \frac{1}{2}\mu^2 \sigma^2, \quad (3.9)$$

where we defined the canonically-normalized field

$$\pi_c^2 \equiv 2M_{\text{pl}}^2|\dot{H}| \pi^2. \quad (3.10)$$

At energies above ρ , the theory is simply that of two weakly-coupled, free fields. Below ρ , the mixing term $\dot{\pi}_c \sigma$ dominates over the kinetic terms $\dot{\pi}_c^2$ and $\dot{\sigma}^2$, and hence determines the low-energy dynamics—i.e. $\dot{\pi}_c \sigma$ becomes the kinetic term of the theory below ρ . The terms $\dot{\pi}_c^2$ and $\dot{\sigma}^2$ are irrelevant operators in this non-relativistic theory and can be ignored at energies below ρ . We can therefore drop the conventional kinetic terms and study the action

$$\mathcal{L} \approx \rho \dot{\pi}_c \sigma - \frac{(\partial_i \pi_c)^2}{2a^2} - \frac{(\partial_i \sigma)^2}{2a^2} - \frac{1}{2}\mu^2 \sigma^2 + \frac{1}{\xi} \left[\dot{\pi}_c^2 - \frac{(\partial_i \pi_c)^2}{a^2} \right] \sigma + \dots, \quad (3.11)$$

where

$$\xi \equiv \frac{2M_{\text{pl}}^2|\dot{H}|}{m^3} = \frac{2}{\rho} (2M_{\text{pl}}^2|\dot{H}|)^{1/2}. \quad (3.12)$$

We note that, at energies below ρ , the two fields π_c and σ are *not* independent degrees of freedom. In particular, at low energies, σ plays the role of the momentum conjugate to π_c ,

$$p_\pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\pi}_c} = \dot{\pi}_c + \rho \sigma \approx \rho \sigma. \quad (3.13)$$

Hence, while the two-field action (3.9) describes two degrees of freedom, one of the degrees of freedom is very massive, $\omega \sim \rho$, and therefore decouples from the low-energy dynamics (this is demonstrated explicitly in Appendix B). The dynamics of the remaining light degree of freedom are governed by the Lagrangian (3.11). For the range of energies $\rho > \omega > \omega_{\text{new}} = \mu^2/\rho$, this describes a single degree of freedom with non-linear dispersion relation $\omega = k^2/\rho$. Below ω_{new} , the ‘mass term’ $\mu^2 \sigma^2$ becomes more important than the kinetic terms and the dispersion relation becomes linear, $\omega = c_s k$, with $c_s \approx \mu/\rho \ll 1$.

¹⁰The theory is non-relativistic at low energies, so we should expect that some of the intuition derived from relativistic field theory will fail.

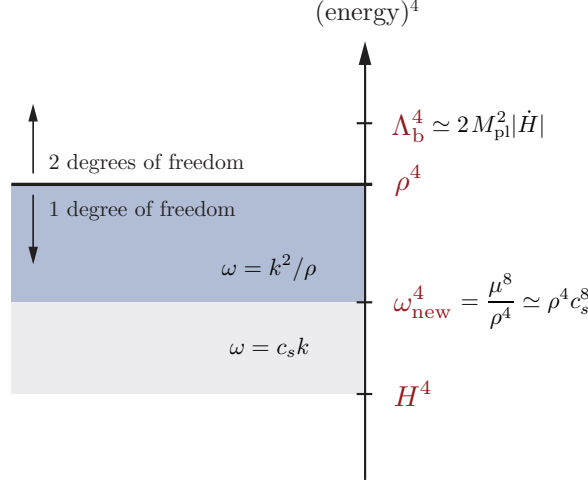


Figure 2: *Relevant energy scales in the π - σ model.*

Symmetry breaking. Since the high-energy behavior of the theory is altered, the symmetry breaking scale will be different in the π - σ model. As we now show, the new symmetry breaking scale will depend on the relative size of ρ^4 and $2M_{\text{pl}}^2|\dot{H}|$.

For $\rho^4 < 2M_{\text{pl}}^2|\dot{H}|$, the symmetry breaking scale is the same as in slow-roll inflation,

$$\Lambda_{\text{b}}^4 = 2M_{\text{pl}}^2|\dot{H}| \quad \text{if} \quad \rho^4 < 2M_{\text{pl}}^2|\dot{H}|. \quad (3.14)$$

This follows simply from the fact that, for energies above ρ , the theory is described by two weakly-interacting fields in a slow-roll background.

For $\rho^4 > 2M_{\text{pl}}^2|\dot{H}|$, determining the symmetry breaking scale requires more care. Below ρ , we work with the effective theory (3.11). The symmetry breaking scale in this theory is modified by the presence of the coupling $\rho\dot{\pi}_c\sigma$. Specifically, the stress tensor derived from (3.11) is $T^{00} = \rho^{1/2}(2M_{\text{pl}}^2|\dot{H}|)^{1/2}\sigma_c + \dots$, where we defined the canonically-normalized field $\sigma_c \equiv \rho^{1/2}\sigma$. Using $[\sigma_c] = [k]^{3/2}$ and the dispersion relation $\omega = k^2/\rho$, we find that the symmetry breaking scale is given by $\Lambda_{\text{b}}^{7/4}\rho^{3/4} = \rho^{1/2}(2M_{\text{pl}}^2|\dot{H}|)^{1/2}$, or

$$\Lambda_{\text{b}}^4 = 2M_{\text{pl}}^2|\dot{H}| \cdot \frac{(2M_{\text{pl}}^2|\dot{H}|)^{1/7}}{\rho^{4/7}} \quad \text{if} \quad \rho^4 > 2M_{\text{pl}}^2|\dot{H}|. \quad (3.15)$$

Weak coupling. Finally, let us determine when our theory is weakly coupled for all energies up to the symmetry breaking scale. To determine the cutoff associated with the Lagrangian (3.11), we define $\tilde{x} = \rho^{1/2}x$, $\tilde{\pi}_c = \rho^{-1/4}\pi_c$ and $\tilde{\sigma}_c = \rho^{-1/4}\sigma$,

$$\mathcal{L} = \dot{\tilde{\pi}}_c\tilde{\sigma}_c - \frac{(\tilde{\partial}_i\tilde{\pi}_c)^2}{2a^2} - \frac{(\tilde{\partial}_i\tilde{\sigma}_c)^2}{2a^2} + \frac{1}{2} \left[\frac{1}{(\Lambda_{\star}^{(1)})^{7/4}} \dot{\tilde{\pi}}_c^2 - \frac{1}{(\Lambda_{\star}^{(2)})^{3/4}} \frac{(\tilde{\partial}_i\tilde{\pi}_c)^2}{a^2} \right] \tilde{\sigma}_c, \quad (3.16)$$

where

$$(\Lambda_{\star}^{(1)})^{7/2} \equiv \frac{2M_{\text{pl}}^2|\dot{H}|}{\rho^{1/2}} \quad \text{and} \quad (\Lambda_{\star}^{(2)})^{3/2} \equiv \frac{2M_{\text{pl}}^2|\dot{H}|}{\rho^{5/2}}. \quad (3.17)$$

The scale $\Lambda_\star^{(1)}$ is always above the symmetry breaking scale (3.15), so the only constraint comes from $\Lambda_\star^{(2)}$. To avoid strong coupling before the symmetry breaking scale, we require $16\pi^2(\Lambda_\star^{(2)})^{3/2} \gg \Lambda_b^{3/2}$, or

$$\rho^4 \ll (16\pi^2)^{7/4} \cdot 2M_{\text{pl}}^2 |\dot{H}|. \quad (3.18)$$

As we explain in Appendix E, this bound is not a unitarity bound, but arises from demanding that the loop expansion is well-defined.

New physics. Where is the scale of new physics $\omega_{\text{new}}^4 = \rho^4 c_s^8$ compared to the would-be strong coupling scale $2\pi\Lambda_\star^4 \simeq 4\pi M_{\text{pl}}^2 |\dot{H}| c_s^5$ of the small c_s -effective theory? From the above upper limit on ρ , cf. Eqn. (3.18), we derive the following bound

$$\omega_{\text{new}}^2 \ll 30 \cdot c_s^{3/2} \cdot \sqrt{2\pi}\Lambda_\star^2 \approx \mathcal{O}(8) \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{-7/4} H^2. \quad (3.19)$$

We remind the reader that the “ \ll ” in (3.19) reflects our requirement of a controlled perturbative expansion at Λ_b . We find that ω_{new}^2 is suppressed relative to Λ_\star^2 by powers of $c_s \ll 1$. This is consistent with the expectation that in weakly-coupled UV-completions the new physics typically enters parametrically below the strong coupling scale. The proximity to the Hubble scale will be relevant in the next section when we discuss the observational consequences.

Finally, we note that, for $\rho \lesssim (2M_{\text{pl}}^2 |\dot{H}|)^{1/4}$, the scale of new physics can even be below the Hubble scale, $\omega_{\text{new}} < H$. In this case, the dispersion relation is non-linear at horizon crossing. At the level of the dispersion relation, the model is then similar to ghost inflation [22]. However, it is not the same as ghost inflation, as can be seen from the scaling dimensions of the fields. If we assign dimensions $[t] = 1$ and $[x] = \frac{1}{2}$, then in our π - σ model $[\pi] = [\sigma] = \frac{3}{4}$, whereas in ghost inflation $[\pi] = \frac{1}{4}$.

3.2.2 Extrinsic Curvature Terms

We have learned that a weakly-coupled UV-completion can arise if the dispersion relation changes at an energy scale below the would-be strong coupling scale in the effective theory for π . Our π - σ model gave an explicit realization of this idea. In this section we will present an alternative UV-completion that is weakly coupled by virtue of the same change in dispersion.

It is well-known that a modified dispersion relation arises in the effective theory of single-field inflation if extrinsic curvature terms are added to the action in unitary gauge [3]. Consider, for instance, the Lagrangian

$$\mathcal{L} = \mathcal{L}_{\text{sr}} + \frac{1}{2}M_2^4(\delta g^{00})^2 - \frac{1}{2}\bar{M}_2^2(\delta K_\mu^\mu)^2, \quad (3.20)$$

where $(\delta K_\mu^\mu)^2$ is a representative extrinsic curvature term. Going to π -gauge and taking the decoupling limit, we find

$$(\delta K_\mu^\mu)^2 \rightarrow \frac{(\partial_i^2 \pi)^2}{a^4} + H \frac{(\partial_i \pi)^2}{a^2} \frac{\partial_j^2 \pi}{a^2} + 2\pi \frac{\partial_i^2}{a^2} \frac{(\partial_j \pi)^2}{a^2}. \quad (3.21)$$

For small c_s —i.e. $M_2^4 \gg M_{\text{pl}}^2 |\dot{H}|$ —the quadratic action for the Goldstone mode becomes

$$\mathcal{L}_2 \approx 2M_2^4 \left[\dot{\pi}^2 - c_s^2 \frac{(\partial_i \pi)^2}{a^2} - \frac{1}{\rho^2} \frac{(\partial_i^2 \pi)^2}{a^2} \right], \quad (3.22)$$

where

$$\rho^2 \equiv \frac{4M_2^4}{\bar{M}_2^2} . \quad (3.23)$$

We see that the dispersion relation changes from linear to non-linear at $\omega_{\text{new}} = \rho c_s^2$. Hence, the strong coupling associated with the $(\delta g^{00})^2$ operator may be avoided if $\omega_{\text{new}}^4 < 2\pi\Lambda_\star^4 = 4\pi M_{\text{pl}}^2 |\dot{H}| c_s^5$, or $\rho^4 < 4\pi M_{\text{pl}}^2 |\dot{H}| c_s^{-3}$.

To conclude that we have a weakly-coupled model at all energies, we must check that any additional strong coupling scales induced by the new interactions in (3.20) parametrically exceed the symmetry breaking scale. First, we note that the change in the dispersion relation implies a new symmetry breaking scale $\bar{\Lambda}_b$. Repeating the analysis of §2.4.1, we find

$$\bar{\Lambda}_b = (4M_2^4)^{2/5} \rho^{-3/5} . \quad (3.24)$$

Next, we apply the treatment of §2.4.2 to the leading interactions in (3.20). The strong coupling

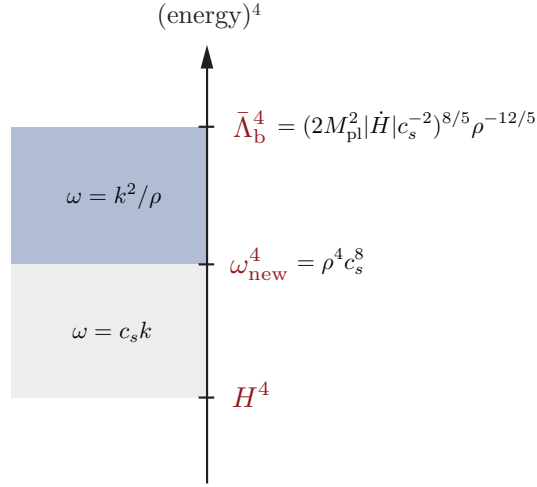


Figure 3: *Relevant energy scales in the model with extrinsic curvature terms.*

scales associated with the operators $\bar{M}_2^2 \dot{\pi} \partial_i^2 (\partial_j \pi)^2$ and $M_2^4 \dot{\pi}^3$ are both larger than $\bar{\Lambda}_b^4$ by a numerical factor. The most stringent bound therefore comes from the operator $M_2^4 \dot{\pi} (\partial_i \pi)^2$, which implies that strong coupling is associated with the following energy scale

$$\bar{\Lambda}_\star = (2M_2^4)^2 \rho^{-7} . \quad (3.25)$$

In Appendix E, we determine the order-one coefficient of the strong coupling scale from the breakdown of perturbative unitarity: $(4\pi)^2 \bar{\Lambda}_\star$. Requiring that $(4\pi)^2 \bar{\Lambda}_\star \geq \bar{\Lambda}_b$ leads to $\rho^4 < (2\pi)^{1/4} 4\pi M_{\text{pl}}^2 |\dot{H}| c_s^{-2}$, or

$$\omega_{\text{new}}^2 \ll \sqrt{c_s} \cdot \sqrt{2\pi} \Lambda_\star^2 \approx \mathcal{O}(6) \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{-5/4} H^2 . \quad (3.26)$$

Again, the scale of new physics is parametrically suppressed (this time only by a factor of $\sqrt{c_s}$) relative to the would-be strong coupling scale of the effective theory with only the $(\delta g^{00})^2$ operator.

4 Observational Consequences

We now set out to answer the question posed in the Introduction: can we detect the ‘new physics’ that is required to enter before the strong coupling scale? Our general strategy will be to compute the three-point functions (or bispectra) generated by our UV-completions and estimate to what extent they can be distinguished from the bispectrum of the strongly-coupled effective theory.

4.1 *Pessimistic View: Intuitive Expectations*

Although the three-point function, in principle, contains a lot of information, in practice, it is difficult to distinguish models whose dominant support is in the same momentum configuration. The signal-to-noise for individual modes is simply too low for a mode-by-mode comparison with the data. Instead, the data is fit to specific templates for the three-point function with fixed momentum dependences. For two different templates whose dominant support lies in the same momentum configuration, the fit with either template will be of similar significance. Two similar templates can therefore only be distinguished if the signal is detected with very high significance using either template.

Our goal is to distinguish the strongly-coupled ‘pure c_s -theory’, $\mathcal{L}_0 = \mathcal{L}_{\text{sr}} + \frac{1}{2}M_2^4(\delta g^{00})^2$, from its weakly-coupled UV-completion. It is easy to see why this is a challenging undertaking. At low energies, the effect of the UV-completion is to add $H^2/\omega_{\text{new}}^2$ -suppressed derivative corrections to the ‘pure c_s -theory’,

$$\mathcal{L} = \mathcal{L}_0 + \Delta\mathcal{L} , \quad (4.1)$$

where $\Delta\mathcal{L} = \mathcal{O}(\frac{H^2}{\omega_{\text{new}}^2})$. It is well-known that \mathcal{L}_0 produces an *equilateral* shape [23]—i.e. the bispectrum peaks in the limit $k_1 = k_2 = k_3$. The challenge is to pick out the subdominant correction to this background. This will only be possible if $\Delta\mathcal{L}$ produces a shape that is significantly non-equilateral and hence distinguishable from the shape produced by \mathcal{L}_0 . Since $\Delta\mathcal{L}$ is produced by higher-derivative corrections one may worry that its bispectrum also peaks in the equilateral configuration and is therefore degenerate with the normalization of \mathcal{L}_0 .

Fortunately, the result of an explicit computation is a bit more optimistic than this. In the next section, we will find that the corrections to the bispectrum arising from generic UV-completions have significant support in the *squashed* configuration, $k_1 = k_2 = \frac{1}{2}k_3$, and are therefore, in principle, distinguishable from the purely equilateral bispectrum. Of course, our calculations are valid only perturbatively in H/ω_{new} , so the corrections are still a subdominant effect in the controlled region.

4.2 *Optimistic View: Bispectra from Higher-Derivative Corrections*

In this section, we describe the change in the bispectrum arising from new physics at ω_{new} . The corrections that arise from new physics generically appear as higher-derivative operators suppressed by the scale ω_{new} . There are only a few operators that could potentially contribute to the three-point function. For example, the corrections to the ‘pure c_s -theory’ arise from terms like $\delta g^{00} \partial^2 \delta g^{00}$ or from extrinsic curvature terms. For concreteness, we will study the corrections in the specific case of the π - σ model; however, one should keep in mind that these corrections

will be generic to most UV-completions. Indeed, the dominant correction to the bispectrum is identical in both the π - σ model and the extrinsic curvature model. Many of the details of the calculation can be found in Appendix C.

From the expression $\omega_{\text{new}} = \mu c_s = \mu^2/\rho$, we see that for fixed c_s the scale of new physics can be changed continuously by altering the parameter μ . Requiring that the theory is weakly coupled at all energy scales implies an upper bound on μ , but there is no lower bound (although small values may require fine-tuning). Unfortunately, an analytic treatment isn't possible for arbitrary values of μ . However, the limits $\mu c_s \gg H$ and $\mu = 0$ are both calculable and we can try to infer the general behavior from these limits.

When $\mu c_s \gg H$, we integrate out σ to get the following effective action for π ,

$$\mathcal{L}_{\text{eff}} = M_{\text{pl}}^2 |\dot{H}| \left[-(\partial_\mu \pi)^2 + \frac{\rho^2}{\mu^2} \left(\dot{\pi} - \frac{1}{2}(\partial_\mu \pi)^2 \right) \left(1 + \frac{\square}{\mu^2} + \dots \right) \left(\dot{\pi} - \frac{1}{2}(\partial_\mu \pi)^2 \right) \right] + \dots \quad (4.2)$$

Dropping $\frac{\square}{\mu^2} + \dots$ reproduces the conventional small speed of sound model \mathcal{L}_0 . Since $\square \sim H^2 c_s^{-2}$ at horizon crossing, all higher-derivative terms may be treated perturbatively as an expansion in $\frac{H^2}{\omega_{\text{new}}^2}$. We will compute the power spectrum and the bispectrum of the comoving curvature perturbation $\zeta = -H\pi$,

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \rangle = (2\pi)^3 P_\zeta(k_1) \delta(\mathbf{k}_1 + \mathbf{k}_2) , \quad (4.3)$$

$$\langle \zeta_{\mathbf{k}_1} \zeta_{\mathbf{k}_2} \zeta_{\mathbf{k}_3} \rangle = (2\pi)^3 B_\zeta(k_1, k_2, k_3) \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) . \quad (4.4)$$

We reproduce the leading power spectrum for the ‘pure c_s -theory’,

$$\Delta_\zeta \equiv k^3 P_\zeta = \frac{1}{4} \frac{H^2}{M_{\text{pl}}^2 c_s \epsilon} . \quad (4.5)$$

By rotational and translational invariance, the bispectrum is only a function of the three magnitudes k_1 , k_2 and k_3 . Moreover, for scale-invariant fluctuations we can extract an overall factor of say k_3^{-6} and write the remaining bispectrum as a function of the rescaled variables

$$x_1 \equiv \frac{k_1}{k_3} \quad \text{and} \quad x_2 \equiv \frac{k_2}{k_3} . \quad (4.6)$$

To compute the bispectrum up to order $\frac{H^2}{\omega_{\text{new}}^2}$, there are three interactions that play a significant role,

$$L_{\text{int}}^{(0)} = \mathcal{C} \cdot \frac{aH}{c_s^2} \cdot \zeta' (\partial_i \zeta)^2 , \quad (4.7)$$

$$L_{\text{int}}^{(1)} = \mathcal{C} \cdot \frac{H^2}{\omega_{\text{new}}^2} \cdot \zeta' \partial_i^2 \zeta' , \quad (4.8)$$

$$L_{\text{int}}^{(2)} = \mathcal{C} \cdot \frac{H^2}{\omega_{\text{new}}^2} \cdot \frac{aH}{c_s^2} \cdot \zeta' \frac{c_s^2 \partial_j^2}{(aH)^2} (\partial_i \zeta)^2 . \quad (4.9)$$

where $L_{\text{int}} \equiv a^4 \mathcal{L}_{\text{int}}$ and $\mathcal{C} \equiv \frac{M_{\text{pl}}^2 |\dot{H}|}{H^4} = (4 \Delta_\zeta c_s)^{-1} \approx \text{const}$. In Appendix C, we compute the bispectra associated with each of these interactions. To make contact with CMB observations, it

proves useful to define an amplitude

$$f_{\text{NL}} \equiv \frac{5}{18} \frac{B_\zeta(1, 1, 1)}{\Delta_\zeta^2} , \quad (4.10)$$

and a shape function [29]

$$S(x_1, x_2) \equiv (x_1 x_2)^2 \cdot \frac{B_\zeta(x_1, x_2, 1)}{B_\zeta(1, 1, 1)} . \quad (4.11)$$

To define the correlation between two distinct shapes S and S' we introduce the scalar product [29]

$$F(S, S') \equiv \int_{\mathcal{V}} S(x_1, x_2) S'(x_1, x_2) \omega(x_1, x_2) dx_1 dx_2 , \quad (4.12)$$

where the integrals are only over physical momenta satisfying the triangle inequality: $0 \leq x_1 \leq 1$ and $1 - x_1 \leq x_2 \leq 1$. We introduced a weight function in the integral, $\omega(x_1, x_2) \equiv (1 + x_1 + x_2)^{-1}$, to achieve that the scalar product exhibits the same scaling as the optimal CMB estimator [30]. As a measure of the degree of correlation between two shapes we use the normalized scalar product or ‘cosine’ [29]

$$\mathcal{C}(S, S') \equiv \frac{F(S, S')}{\sqrt{F(S, S)F(S', S')}} . \quad (4.13)$$

The leading non-Gaussianity is generated by $L_{\text{int}}^{(0)}$, the interaction associated with the operator $(\delta g^{00})^2$. Computing the shape of the bispectrum in the in-in formalism (see Appendix C) gives [23]

$$S^{(0)} = -\frac{9}{17} \cdot \frac{X_1^6 - 3X_1^4 X_2^2 + 11X_1^3 X_3^3 - 4X_1^2 X_2^4 - 4X_1 X_2^2 X_3^3 + 12X_3^6}{X_3^3 X_1^3} , \quad (4.14)$$

where

$$X_1 = 1 + x_1 + x_2 , \quad (4.15)$$

$$X_2 = (x_1 x_2 + x_1 + x_2)^{1/2} , \quad (4.16)$$

$$X_3 = (x_1 x_2)^{1/3} . \quad (4.17)$$

The shape $S^{(0)}$ peaks in the equilateral configuration $x_1 = x_2 = 1$. In fact, we will use $S^{(0)}$ as our definition of the ‘equilateral shape’ $S_{\text{equil}} \equiv S^{(0)}$. The amplitude of this equilateral non-Gaussianity is [23]

$$f_{\text{NL}}^{\text{equil.}} \equiv f_{\text{NL}}^{(0)} = \frac{85}{324} \left(1 - \frac{1}{c_s^2} \right) \simeq -\frac{1}{4c_s^2} . \quad (4.18)$$

A second shape of interest arises as a special linear combination of the shapes associated with the operator $M_2^4 (\delta g^{00})^2$ and the operator $M_3^4 (\delta g^{00})^3 \rightarrow \tilde{c}_3 \cdot \mathcal{C} \cdot \frac{aH}{c_s^4} (\zeta')^3$. Both operators individually produce similar (but not identical) equilateral shapes. By tuning the coefficient \tilde{c}_3 , we define an ‘orthogonal shape’ [23] by the value of \tilde{c}_3 for which the correlation between the equilateral shape $S^{(0)} \equiv S_{\text{equil}}$ and the orthogonal shape S_{ortho} vanishes

$$\mathcal{C}(S_{\text{equil}}, S_{\text{ortho}}) \equiv 0 . \quad (4.19)$$

This occurs when $\tilde{c}_3 \simeq -5.4$.

Both $L_{\text{int}}^{(1)}$ and $L_{\text{int}}^{(2)}$ contribute to the bispectrum at leading order in $\frac{H^2}{\omega_{\text{new}}^2}$. The interaction $L_{\text{int}}^{(1)}$ furthermore leads to a small correction to the power spectrum (see Appendix C). However, this correction is scale-invariant and so only corresponds to an unobservable shift in the amplitude. One may hope that the shape of the correction to the bispectrum, $S^{(1)}$, leaves a more detectable imprint. However, we find that $S^{(1)}$ is not significantly different from the equilateral shape $S^{(0)}$. In fact, the cosine between the shape $S^{(1)}$ and the orthogonal shape is only

$$\mathcal{C}(S^{(1)}, S_{\text{ortho}}) = 0.21 . \quad (4.20)$$

As a result, the correction to the bispectrum from $L_{\text{int}}^{(1)}$ will, in practice, be difficult to distinguish from a shift in the value of M_2 .

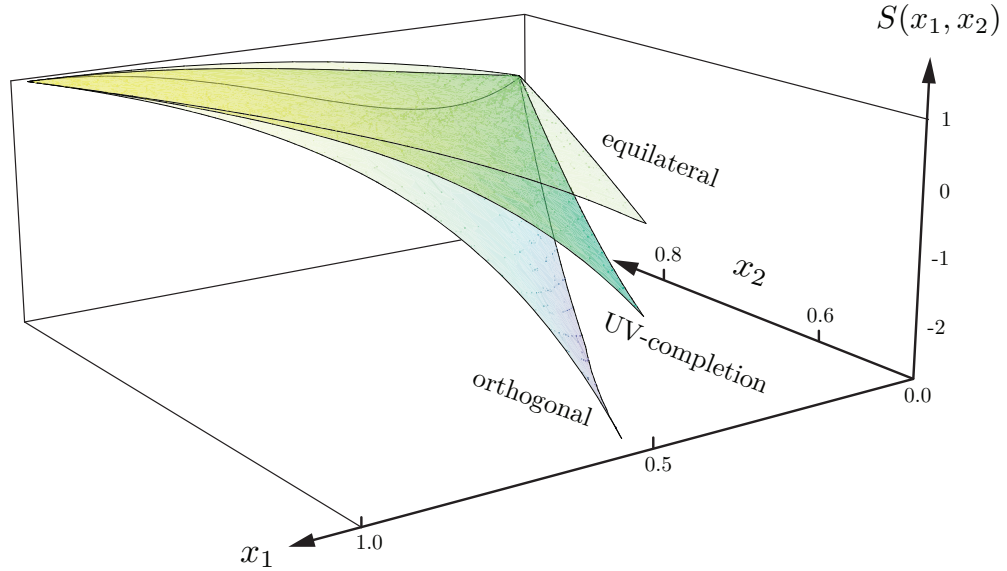


Figure 4: Comparison between the shape produced by the dominant higher-derivative correction in both UV-completions, the equilateral shape and the orthogonal shape. All shapes are normalized relative to the amplitude in the equilateral configuration.

The more interesting correction comes from $L_{\text{int}}^{(2)}$. This is the same interaction as appears in the extrinsic curvature term described in §3.2.2. The shape of the corresponding contribution to the bispectrum is

$$S^{(2)} = \frac{27}{13} \cdot \frac{X_1^8 - 3X_1^6 X_2^2 - 7X_1^4 X_2^4 + 12X_1^2 X_2^6 + 17X_1^5 X_3^3}{X_3^3 X_1^5} \times \frac{-43X_1^3 X_2^2 X_3^3 + 36X_1 X_2^4 X_3^3 + 66X_1^2 X_3^6 - 48X_2^2 X_3^6}{X_3^3 X_1^5} . \quad (4.21)$$

The shape $S^{(2)}$ shows a significant peak in the squashed configuration (see Figure 4).¹¹ This is reflected in the cosine between this shape and the orthogonal shape

$$\mathcal{C}(S^{(2)}, S_{\text{ortho}}) = 0.56 . \quad (4.22)$$

¹¹The same shape also appears in the effective theory of single-field inflation if certain extrinsic curvature operators are considered [31].

This overlap with the orthogonal shape gives us hope that $S^{(1)}$ could be detectable in a measurement using the orthogonal template. However, although quite distinct in shape, this correction will be very hard to extract from the data if the amplitude of the correction is small. This is required by our wish to maintain perturbative control over our calculations, but does not have to be the case more generally.

4.3 Realistic View: Observational Prospects

In the previous section, we computed the correction to the bispectrum produced by physics near the Hubble scale. We found that the correction $S^{(2)}$ has a significant overlap with the orthogonal shape, which by definition has zero overlap with the equilateral shape $S^{(0)}$. When we measure non-Gaussianity using the orthogonal template, we are therefore effectively projecting out the contribution to the signal from $S^{(0)}$. In the ‘pure c_s -theory’, the signal is only coming from $S^{(0)}$, so we don’t expect to measure $f_{\text{NL}}^{\text{ortho.}}$ in that case. In contrast, in our UV-completions, we can get contributions both to $f_{\text{NL}}^{\text{equil.}}$ (mostly from $S^{(0)}$) and to $f_{\text{NL}}^{\text{ortho.}}$ (mostly from $S^{(2)}$). We propose to use this correlated signature as a diagnostic for our UV-completions. In this section, we will use this fact to give a rough estimate for when the contribution $S^{(2)}$ can be measured in future experiments.

The shape computed in (4.21) is the first term in a perturbative expansion in $\frac{H^2}{\omega_{\text{new}}^2}$. Therefore, the amplitude of this contribution is suppressed by $\frac{H^2}{\omega_{\text{new}}^2} < 1$. From the calculation in Appendix C, we infer that the relative size of the two bispectra in the equilateral limit is

$$B_{\zeta}^{(2)}(1, 1, 1) = \frac{26}{51} \cdot \frac{H^2}{\omega_{\text{new}}^2} \cdot B_{\zeta}^{(0)}(1, 1, 1) . \quad (4.23)$$

Using the standard definition of f_{NL} (4.10), we find that the relative contribution to f_{NL} coming from the higher-derivative correction is given by

$$f_{\text{NL}}^{(2)} \simeq \frac{1}{2} \frac{H^2}{\omega_{\text{new}}^2} f_{\text{NL}}^{(0)} . \quad (4.24)$$

In practice, we do not measure this f_{NL} , but rather use the equilateral and orthogonal templates to determine $f_{\text{NL}}^{\text{equil.}}$ and $f_{\text{NL}}^{\text{ortho.}}$. The contribution from (4.24) to $f_{\text{NL}}^{\text{equil.}}$ can be absorbed by a change in the normalization of $L_{\text{int}}^{(0)}$ (the size of c_s).¹² We will therefore focus on $f_{\text{NL}}^{\text{ortho.}}$.

The definition of the cosine in (4.13) was chosen to related the experimental bounds from different templates [29]. This allows us to estimate the contribution from (4.24) to $f_{\text{NL}}^{\text{ortho.}}$ as

$$\Delta f_{\text{NL}}^{\text{ortho.}} \simeq \frac{1}{2} \mathcal{C}(S^{(2)}, S_{\text{ortho}}) \cdot \sqrt{\frac{S_{\text{ortho.}} \cdot S_{\text{ortho.}}}{S_{\text{equil.}} \cdot S_{\text{equil.}}}} \cdot \frac{H^2}{\omega_{\text{new}}^2} f_{\text{NL}}^{\text{equil.}} . \quad (4.25)$$

Using $\mathcal{C}(S^{(2)}, S_{\text{ortho}}) = 0.56$ and $S_{\text{ortho}} \cdot S_{\text{ortho}} \simeq \frac{9}{4} S_{\text{equil.}} \cdot S_{\text{equil.}}$, we get

$$\Delta f_{\text{NL}}^{\text{ortho.}} \sim \frac{1}{2} \frac{H^2}{\omega_{\text{new}}^2} f_{\text{NL}}^{\text{equil.}} . \quad (4.26)$$

¹²In the perturbative regime we will use $f_{\text{NL}}^{\text{equil.}} \simeq f_{\text{NL}}^{(0)}$.

This is the size of the predicted non-Gaussianity measured with the orthogonal template if the non-Gaussianity measured using the equilateral template is $f_{\text{NL}}^{\text{equil}}$.

As we explained before—cf. (3.19)—the ratio $H^2/\omega_{\text{new}}^2$ satisfies

$$\frac{1}{8} \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{7/4} \ll \frac{H^2}{\omega_{\text{new}}^2} < 1. \quad (4.27)$$

Combining (4.26) and (4.27), we find

$$6 \cdot \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right)^{11/4} \ll \Delta f_{\text{NL}}^{\text{ortho.}} < 50 \cdot \left(\frac{f_{\text{NL}}^{\text{equil.}}}{100} \right). \quad (4.28)$$

Interestingly, we get a lower bound on the expected signal from the requirement that the theory be perturbative at Λ_{b} —which is responsible for the “ \ll ” in (4.27). However, we emphasize that we are pushing the validity of the perturbative calculation as we approach the upper limit in (4.28). For the contribution to the orthogonal shape to be detectable in the regime of perturbative control (i.e. $\Delta f_{\text{NL}}^{\text{ortho.}} \gtrsim 10$ with $H < \omega_{\text{new}}$), we require a detection of equilateral non-Gaussianity near its current upper limit, $|f_{\text{NL}}^{\text{equil.}}| \lesssim 250$ [28]. Even in this optimistic case, the signal will be hard to detect with future CMB experiments.¹³

The above conclusions relied on limiting ourselves to the perturbative regime, $H \ll \omega_{\text{new}}$. Given the strong upper bound on the scale of new physics, the more likely scenario is when $\omega_{\text{new}} \sim H$, where our perturbative techniques break down. The perturbative calculation may be suggestive that this limit could generate a significant contribution to $f_{\text{NL}}^{\text{ortho.}}$. This is a well-motivated scenario that would become relevant in the event of a detection of equilateral non-Gaussianity by the Planck satellite.

4.4 The $\omega_{\text{new}} \rightarrow 0$ Limit

The calculation of the previous section suggests that a significant orthogonal component could be generated as $\omega_{\text{new}} = \mu^2/\rho \rightarrow H$. Unfortunately, we were unable to explicitly calculate the bispectrum in this limit. In the absence of an analytic bispectrum calculation for all values of μ , it is still interesting to study the extreme limit, $\mu \rightarrow 0$. This limit is also interesting simply because it is a new single-field model whose dynamics haven’t previously been considered in the literature. In this section, we will determine the observational signatures of the $\mu \rightarrow 0$ limit of the π - σ model, with many details left to Appendix D.

The Lagrangian describing this limit is

$$\mathcal{L} \approx \rho \dot{\pi}_c \sigma - \frac{1}{2} \frac{(\partial_i \pi_c)^2}{a^2} - \frac{1}{2} \frac{(\partial_i \sigma)^2}{a^2} - \underbrace{\frac{1}{\xi} (\partial_\mu \pi_c)^2 \sigma}_{\mathcal{L}_{\text{int}}}, \quad (4.29)$$

where $\xi \equiv \frac{2M_{\text{pl}}^2 |\dot{H}|}{m^3} = \frac{2}{\rho} (2M_{\text{pl}}^2 |\dot{H}|)^{1/2}$. The interaction \mathcal{L}_{int} will be responsible for generating a measurable bispectrum. Focusing first on the quadratic action, we get the following equations of

¹³To get a sense for the significant observational challenge that this implies, we remind the reader of the WMAP 95% C.L. constraints [28]: $-214 < f_{\text{NL}}^{\text{equil.}} < 266$ and $-410 < f_{\text{NL}}^{\text{ortho.}} < 6$. Forecasted 1- σ errors for future experiments are [32, 33]: $\sigma(f_{\text{NL}}^{\text{equil.}}, f_{\text{NL}}^{\text{ortho.}}) \sim 30$ (Planck [34]) and $\sigma(f_{\text{NL}}^{\text{equil.}}, f_{\text{NL}}^{\text{ortho.}}) \sim 10$ (CMBPol [35], CORe [36]).

motion

$$\ddot{\pi}_c + 5H\dot{\pi}_c + \frac{k^4}{\rho^2 a^4} \pi_c = 0 \quad \text{and} \quad \frac{k^2}{a^2} \sigma = \rho \dot{\pi}_c . \quad (4.30)$$

Note the unusual Hubble friction factor of $5H$ (rather than $3H$) that occurs in this model. This is a consequence of σ (rather than $\dot{\pi}_c$) being the canonical momentum of π_c . Both of these facts are important for achieving a scale-invariant power spectrum.

The mode functions are most easily determined in conformal time. As usual, we will quantize the system by writing $\hat{\pi}_c(\mathbf{k}, \tau) = \pi_k(\tau) \hat{a}_{\mathbf{k}} + \pi_k^*(\tau) \hat{a}_{-\mathbf{k}}^\dagger$, where $\pi_k(\tau)$ is a solution to the equations of motion. Being careful to define the Bunch-Davies vacuum when σ is the canonical momentum (see Appendix D), one finds

$$\pi_k(\tau) = (H\tau)^2 \frac{\sqrt{-k^2\tau}}{\rho} H_{5/4}^{(1)}\left(\frac{1}{2} \frac{H}{\rho} (k\tau)^2\right) , \quad (4.31)$$

where $H_\nu^{(1)}(x)$ is the Hankel function of the first kind. The resulting power spectrum for ζ is scale-invariant, with amplitude

$$\Delta_\zeta \equiv k^3 |\zeta_k^{(o)}|^2 \sim \frac{H^2}{M_{\text{pl}}^2 \epsilon} \left(\frac{\rho}{H}\right)^{1/2} . \quad (4.32)$$

The bispectrum calculation proceeds as usual if we use the interaction $\mathcal{L}_{\text{int}} = -\frac{1}{\xi} (\partial_\mu \pi_c)^2 \sigma$, with $\sigma = -\frac{\rho}{H\tau k^2} \pi_c'$. Unfortunately, the bispectrum can only be computed numerically. The resulting shape is very similar to the one generated by the $M_2^4 \dot{\pi} (\partial_\mu \pi)^2$ interaction in the ‘pure c_s -theory’,

$$\mathcal{C}(S_{\mu=0}, S^{(0)}) = 0.99 . \quad (4.33)$$

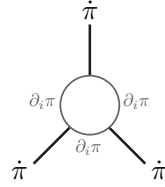
The similarity of the shapes can be understood at the level of the interaction Hamiltonian. Because $\sigma \propto \dot{\pi}/k^2$, the basic kinematic momentum factor is still $(\mathbf{k}_1 \cdot \mathbf{k}_2 + \text{perms.})$, arising from the $\partial_i \pi \partial^i \pi$ -part of \mathcal{L}_{int} . Moreover, although the mode functions are different, they are still functions of k that vanish exponentially inside the horizon. In contrast, the corrections computed in the previous section included interactions of the form $\dot{\pi} \partial_i \partial_j \pi \partial^i \partial^j \pi$. The resulting kinematic factor therefore differs and introduces a significant signal in the squashed momentum configuration.

The limit $\mu \rightarrow 0$ ($\omega_{\text{new}} \rightarrow 0$) in the π - σ model has an analogue in the extrinsic curvature model discussed in §3.2.2. In that case, the limit $\omega_{\text{new}} \rightarrow 0$ corresponds to the limit $\dot{H} \rightarrow 0$ with finite M_2 and \bar{M}_2 . This model is typically referred to as “ghost inflation” [22]. The predictions of ghost inflation are discussed in detail in [23]. Its correlation with the orthogonal shape is small, $\mathcal{C}(S_{\text{ghost}}, S_{\text{ortho}}) = 0.25$, making it difficult to distinguish the bispectrum of ghost inflation from the equilateral bispectrum of the ‘pure c_s -theory’.

5 Comments on Naturalness

In this paper we have considered *weak coupling* as a new criterium to narrow down the in principle vast space of effective theories of inflation [3, 5]. So far, we have not required our theories to be natural. This is in the same spirit as the analogous situation in the Standard Model, where an unnaturally light Higgs particle is introduced to keep the theory of massive gauge bosons weakly coupled. Physics beyond the Standard Model is required to explain the small Higgs mass. Similarly, additional structures in the high-energy theory may be required to make our theories natural. We implicitly assumed that this doesn't change the low-energy phenomenology at $\omega < \omega_{\text{new}}$. In the effective theory of inflation [3, 5, 37, 23] *naturalness* was proposed as a basic criterium to focus on the interesting regimes of the parameter space of couplings. In this section, we comment briefly on how this notion of naturalness may be modified in our weakly-coupled UV-completions.

Let us first review the argument of [23] concerning the natural values of parameters in the effective theory with small c_s . Our starting point is the action (2.1), with $M_2^4 \approx \frac{1}{2} M_{\text{pl}}^2 |\dot{H}| c_s^{-2}$ and a priori unknown values for the coefficients M_n^4 for $n > 2$. The interaction $M_3^4 (\delta g^{00})^3 = M_3^4 (\dot{\pi}^3 + \dots)$ will be generated from $M_2^4 (\delta g^{00})^2 = M_2^4 (\dot{\pi}^2 + \dot{\pi} \frac{(\partial_i \pi)^2}{a^2} + \dots)$ via the following loop



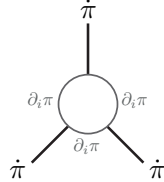
$$= \dot{\pi}^3 \int d\omega d^3k \frac{k^6}{(\omega^2 - c_s^2 k^2)^3} = \frac{\Lambda_{\text{uv}}^4}{c_s^9} \dot{\pi}^3 \equiv M_3^4 \dot{\pi}^3, \quad (5.1)$$

where Λ_{uv} is the UV-cutoff of the loop integral. If there is *no* new physics before the strong coupling scale Λ_* , the UV-cutoff is at least Λ_* . In this case, loop effects generate a large value for M_3 ,

$$\Lambda_{\text{uv}}^4 \rightarrow \Lambda_*^4 = M_2^4 c_s^7 \Rightarrow M_3^4 \sim \frac{M_2^4}{c_s^2}. \quad (5.2)$$

As we have seen above, such a large value for M_3 can be of observational relevance since it is a prerequisite for generating an orthogonal shape for the bispectrum: $M_3^4 \approx -5.4 M_2^4 c_s^{-2}$ [23].

This renormalization argument is modified in our weakly-coupled examples. Both of our UV-completions are characterized by a change in the dispersion relation from $\omega = c_s k$ to $\omega = k^2/\rho$ at $\omega_{\text{new}} = c_s^2 \rho$. This changes the high-energy behavior of the loop generating M_3 . With the new dispersion relation, the loop integral in (5.1) scales as



$$= \dot{\pi}^3 \int_{\Lambda} d\omega d^3k \frac{k^6}{(\omega^2 - \frac{k^4}{\rho^2})^3} \sim \frac{\rho^{9/2}}{\Lambda^{1/2}} \dot{\pi}^3 = \frac{1}{c_s^9} \frac{\omega_{\text{new}}^{9/2}}{\Lambda^{1/2}} \dot{\pi}^3. \quad (5.3)$$

The integral is now IR-dominated and hence converges in the UV. The effective UV-cutoff of the low-energy theory becomes $\Lambda_{\text{uv}}^4 \rightarrow \omega_{\text{new}}^4$. In weakly-coupled UV-completions we expect the scale of new physics to be parametrically below the scale of strong coupling. In both of our examples

we have seen this expectation confirmed: In the π - σ -model we found $\omega_{\text{new}}^4 < c_s^3 \Lambda_\star^4$, while the model with extrinsic curvature requires $\omega_{\text{new}}^4 < c_s \Lambda_\star^4$. This suppresses the loop-generated size of M_3 ,

$$M_3^4 \sim \frac{\omega_{\text{new}}^4}{c_s^9} < \frac{M_2^4}{c_s^2}. \quad (5.4)$$

A change in the dispersion relation below the strong coupling scale therefore naively seems to be a simple way to realize both weak coupling and a naturally small value of M_3 . However, what happens above the scale ω_{new} in our specific examples is more complicated than just a change in the dispersion relation. The physics responsible for changing the dispersion relation may also change the interaction Hamiltonian or even the scaling properties of the theory. Therefore, we must consider the full UV-completion to determine the natural values of the low-energy parameters within a given model.

In any case, it should be clear that in our UV-completions there is nothing special about the scale Λ_\star . As a result, any divergent integrals will be cut off at Λ_b . While some divergences are regulated by a change in the dispersion relation, both models contain additional operators that were fine-tuned. This is most clear in the π - σ model where we introduced a scalar σ with a mass term $\mu^2 \sigma^2$. Corrections to μ^2 are divergent and suggest that $\mu \sim \Lambda_b$ in a natural theory. Furthermore, one easily generates additional operators like $\tilde{\mu} \sigma^3$. In this model, the question of fine-tuning has been moved to a scalar potential, but has not been resolved. One would suspect that supersymmetry would resolve the problem, but constructing supersymmetric completions of our theories lies beyond the scope of this work.

6 Discussion

With the increasing precision of CMB measurements by Planck and other experiments on the horizon, it is timely to ask what we can hope to learn about inflation from these observations. In the absence of measurable B-mode polarization or deviations from Gaussianity, the amount of information is limited to the amplitude and the scale-dependence of the power spectrum of primordial density fluctuations. In contrast, if non-Gaussianity were detected, it would provide an entire function worth of information about the physics of inflation [8].

A powerful way to describe non-Gaussianity in single-field models is the effective theory of inflation [3]. In this approach, large interactions are associated with a small sound speed. The dominant signals arise from two distinct operators both of which produce equilateral shapes of non-Gaussianity. By fine-tuning the relative coefficients of these two operators one creates the so-called orthogonal shape [23]. In this paper, we discussed the physical implications of a detection of non-Gaussianity with approximately equilateral shape. We showed that the associated effective theories become strongly coupled far below the symmetry breaking scale, but not far above the Hubble scale. We compared this situation to the Standard Model of particle physics, where WW scattering becomes strongly coupled around the TeV scale, unless ‘new physics’, such as a light Higgs particle, is introduced below the would-be strong coupling scale. Similarly, for the effective theory of inflation to be weakly coupled at all energies, we require new physics to appear below the strong coupling scale. Since the scale of new physics can’t be too far above the

Hubble scale, one may hope that it doesn't completely decouple from measurements of primordial non-Gaussianity. We computed the signatures of candidate UV-completions perturbatively in the small ratio $H^2/\omega_{\text{new}}^2$. Interestingly, we found that the leading corrections to the bispectrum have a significant contribution in the squashed limit, and can therefore, in principle, be distinguished from the purely equilateral signal. In practice, detecting the new physics requires a high-significance detection of the dominant equilateral signal and assumes that $H^2/\omega_{\text{new}}^2$ is not too small. Since observational constraints put a rather strong lower bound on $H^2/\omega_{\text{new}}^2$, even the minimal correction to the equilateral shape may not be totally out of reach of future experiments.

A number of future studies suggest themselves:

- Although our analytic computations were performed for $H^2/\omega_{\text{new}}^2 \ll 1$, it is more natural for $H^2/\omega_{\text{new}}^2$ to be of order one. Computing the shape of the three-point function in this case would be particularly interesting, but would require a non-perturbative treatment. Using our perturbative calculations as a guide, we may suspect that the bispectrum for $\omega_{\text{new}} \sim H$ could deviate significantly from the equilateral shape. Checking this intuition explicitly will likely require numerical work. We plan to return to this question in the future.
- By demanding that inflationary models with small sound speed are weakly coupled at all energies, we were lead to consider fairly novel effective theories. However, there was nothing special about considering small sound speed. In fact, new physics near the Hubble scale may appear in many other models with large non-Gaussianities—e.g. the multi-field effective theory [5] contains a large number of additional interactions, many of which lead to strong coupling near the Hubble scale. It would be interesting to extend our analysis to these cases and investigate the observational signatures of their UV-completions.
- Our conclusions aren't restricted to the bispectrum. For example, a large trispectrum generated by the operator $M_4^4(g^{00} + 1)^4$ [37] also implies a strong coupling scale close to the Hubble scale. Again, if new physics becomes important near the Hubble scale, it may not decouple from observations. The nature of the UV-completions that give rise to these large values of M_4 may differ from those relevant to small c_s , warranting a separate analysis.
- We have not required our theories to be technically natural. Constructing technically natural versions of our UV-completions could lead to additional structures at high energies. Again, this is similar to the situation in the Standard Model where an unnaturally light Higgs is made natural by supersymmetry. It is interesting to note that supersymmetric models of inflation generically include particles with Hubble scale masses. Hence, supersymmetry could potentially offer a compelling explanation of 'new physics' at the Hubble scale. It would be worthwhile to develop supersymmetric versions of our UV-completions to either confirm that the additional physics decouples or to explore its low-energy signatures.

Acknowledgements

We are grateful to Nathaniel Craig, Thomas Dumitrescu, Raphael Flauger, Zohar Komargodski, Enrico Pajer, Rafael Porto, Soo-Jong Rey, Leonardo Senatore, Amit Yadav and Matias Zaldarriaga for helpful discussions. The research of D.B. is supported by the National Science

Foundation under PHY-0855425 and a William D. Loughlin Fellowship at the Institute for Advanced Study. D.B. thanks the Institute for Theoretical Physics, Madrid and the theory group at Cornell for hospitality. The research of D.G. is supported by the DOE under grant number DE-FG02-90ER40542 and the Martin A. and Helen Chooljian Membership at the Institute for Advanced Study. D.G. thanks the Stanford Institute for Theoretical Physics for hospitality. D.B. and D.G. thank the Centro de Ciencias de Benasque Pedro Pascual for hospitality while this work was being completed.

A Strong Coupling in DBI Inflation

In this paper, we studied the effective theory of *fluctuations* around quasi-de Sitter backgrounds. We showed that in the limit of small sound speed these fluctuations become strongly coupled not far from the energy scale associated with the cosmological experiment, the Hubble scale. In this appendix, we relate these findings to the well-known special properties of the effective theory of the *background* [10, 11].

Inducing a non-trivial sound speed for the fluctuations requires higher-derivative terms to be dynamically important during inflation,

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu\phi)^2 \left(1 + \sum_{n=1}^{\infty} \frac{(\partial_\mu\phi)^{2n}}{M^{4n}} \right). \quad (\text{A.1})$$

In fact, the limit $c_s \ll 1$ requires $(\partial_\mu\phi)^2 \sim M^4$ [10], indicating a breakdown of the standard derivative expansion of the effective theory for the background. To make sense of the derivative expansion in (A.1) seems to require a UV-completion. A famous example for the UV-completion of models with small c_s is Dirac-Born-Infeld (DBI) inflation [12],

$$\mathcal{L} = -M^4 \sqrt{1 - \frac{(\partial_\mu\phi)^2}{M^4}}. \quad (\text{A.2})$$

Although all terms in the expansion in powers of $(\partial_\mu\phi)^2$ are equally important in the limit $c_s \ll 1$, the higher-dimensional boost symmetry of the DBI action nevertheless controls the theory [38]. DBI inflation therefore is an example where the effective theory of the small c_s -background is well-defined.

What about the corresponding effective theory for the fluctuations? What determines the strong coupling scale that we identified in the main text? Because the strong coupling scale is expected to be above the Hubble scale, we can work in the Minkowski limit. Consider the DBI action (A.2). Any background of the form $\dot{\phi} = \text{const.}$ is a solution to the equations of motion. We expand the action in fluctuations around the background, $\phi = \bar{\phi}(t) + \varphi(t, \mathbf{x})$, and choose $\dot{\phi} = M^2 \sqrt{1 - c_s^2}$. The action for the fluctuations is

$$\mathcal{L} = -M^4 c_s \sqrt{1 - 2(1 - c_s^2)^{1/2} \frac{\dot{\varphi}}{c_s^2 M^2} - \frac{(\partial_\mu\varphi)^2}{c_s^2 M^4}}. \quad (\text{A.3})$$

Expanding the square root, we find the quadratic action

$$\mathcal{L}_2 = \frac{1}{2c_s^3}(\dot{\varphi}^2 - c_s^2(\partial_i\varphi)^2). \quad (\text{A.4})$$

For sufficiently large $\dot{\phi}$, a small sound speed, $c_s \ll 1$, is generated for the fluctuations. However, the expansion of the square root was only valid when $\dot{\varphi}^2 < M^4 c_s^4 (1 - c_s^2)^{-1}$. Moreover, we were justified in treating φ as the fluctuation around a background as long as $\dot{\varphi}^2 < M^4$. Hence, the theory for the fluctuations is strongly coupled when

$$M^4 c_s^4 (1 - c_s^2)^{-1} < \dot{\varphi}^2 < M^4. \quad (\text{A.5})$$

This range of energies is consistent with our previous analysis. In particular, as before, the strong coupling scale is suppressed by four powers of c_s relative to the symmetry breaking scale.

B Dynamics of the π - σ Model

In this appendix, we describe the dynamics of the π - σ model. Our starting point is the quadratic action

$$\mathcal{L}_2 = -\frac{1}{2}(\partial_\mu \pi_c)^2 - \frac{1}{2}(\partial_\mu \sigma)^2 + \rho \dot{\pi}_c \sigma - \frac{1}{2}\mu^2 \sigma^2 . \quad (\text{B.1})$$

The corresponding equations of motion are

$$\ddot{\pi} + 3H\dot{\pi} + k_p^2 \pi = -\rho[3H\sigma + \dot{\sigma}] , \quad (\text{B.2})$$

$$\ddot{\sigma} + 3H\dot{\sigma} + k_p^2 \sigma = -\mu^2 \sigma + \rho \dot{\pi} , \quad (\text{B.3})$$

where $k_p = k/a(t)$ is the physical momentum. Since we will always be working in the limit $\rho \gg H$, terms proportional to ρ become important well inside the horizon. The dynamics in the regime $k \sim \rho$ are therefore well-approximated by the flat space limit¹⁴ $H \rightarrow 0$. In this limit, we can find the mode solutions exactly using the ansatz $\pi = Ae^{i\omega t}$ and $\sigma = Be^{i\omega t}$. The equations of motion become algebraic equations for A , B and ω ,

$$[\omega^2 - k^2] \pi = \rho(i\omega) \sigma , \quad (\text{B.4})$$

$$[\omega^2 - (k^2 + \mu^2)] \sigma = -i\rho\omega \pi . \quad (\text{B.5})$$

Combining these two equations, we find

$$\omega_\pm^2 = k^2 + \frac{\rho^2 + \mu^2}{2} \pm \sqrt{\frac{(\rho^2 + \mu^2)^2}{4} + \rho^2 k^2} . \quad (\text{B.6})$$

When $k \gg \rho > \mu$, we have two positive frequency modes with $\omega \sim k$. This is not surprising. At very high energies π and σ are essentially independent free fields. When $k \ll \rho$, we expand the square root in (B.6) to find

$$\omega_\pm^2 \simeq k^2 + \frac{\rho^2 + \mu^2}{2} \pm \left(\frac{\rho^2 + \mu^2}{2} + \frac{\rho^2}{\rho^2 + \mu^2} k^2 - \frac{\rho^4}{(\rho^2 + \mu^2)^3} k^4 + \dots \right) . \quad (\text{B.7})$$

The ω_+ solution gives rise to a positive frequency mode with $\omega \sim \rho$. This describes a very massive degree of freedom as its energy is always of order ρ . In contrast, the positive ω_- solution corresponds to a positive frequency mode with $\omega \ll \rho$, i.e. a light degree of freedom. The dispersion relation for this mode is

$$\omega_-^2 \simeq \left(1 - \frac{\rho^2}{\rho^2 + \mu^2} \right) k^2 + \frac{\rho^4}{(\mu^2 + \rho^2)^3} k^4 \equiv c_s^2 k^2 + \frac{k^4}{\tilde{\rho}^2} , \quad (\text{B.8})$$

where $c_s^2 = \mu^2/(\rho^2 + \mu^2)$ and $\tilde{\rho}^2 = (\rho^2 + \mu^2)^3/\rho^4$.

We are interested in the behavior for $\mu^2 \ll \rho^2$. In this case, $c_s^2 \approx \mu^2/\rho^2 \ll 1$ and $\tilde{\rho} \approx \rho$. When $\rho > k > c_s \rho$, we see that the second term in (B.8) dominates, and the dispersion relation is $\omega \approx k^2/\rho$. Therefore, for the range of energies $\rho > \omega > c_s^2 \rho$ the mode is effectively described by the free Schrödinger equation. When $\omega < c_s^2 \rho \equiv \omega_{\text{new}}$ (or $k < \mu$), we return to a linear dispersion relation with a small speed of sound, $\omega \approx c_s k$.

¹⁴Including the effects of finite H using a WKB-like approximation is straightforward, but doesn't qualitatively affect our arguments.

As we have seen, when $\omega < \rho$, there is only one degree of freedom. In this limit, the full quadratic action is unnecessary as it includes the heavier mode. The action that describes just the light mode is

$$\mathcal{L}_2 \approx \rho \dot{\pi}_c \sigma - \frac{1}{2} \frac{(\partial_i \pi_c)^2}{a^2} - \frac{1}{2} \frac{(\partial_i \sigma)^2}{a^2} - \frac{1}{2} \mu^2 \sigma^2 . \quad (\text{B.9})$$

Here we have dropped the relativistic kinetic terms, $\dot{\pi}_c^2$ and $\dot{\sigma}^2$, as they introduce corrections which at low energies are suppressed by ω/ρ . The equations of motion now are,

$$k_{\text{p}}^2 \pi = -\rho [3H\sigma + \dot{\sigma}] , \quad (\text{B.10})$$

$$k_{\text{p}}^2 \sigma = -\mu^2 \sigma + \rho \dot{\pi} . \quad (\text{B.11})$$

Repeating our analysis in the $H \rightarrow 0$ limit, we find that

$$\omega^2 = \frac{\mu^2}{\rho^2} k^2 + \frac{k^4}{\rho^2} . \quad (\text{B.12})$$

When $\mu \ll \rho$, this reproduces the dispersion relation of the full quadratic action (B.8).

C Corrections to Vanilla Sound Speed Models

In this appendix, we describe in more detail the predictions of our two-field UV-completion of small speed of sound theories. We will consider the limit $\omega_{\text{new}} \equiv \mu^2/\rho > H$, where the corrections to the ‘pure c_s -theory’, $\mathcal{L}_0 \equiv \mathcal{L}_{\text{sr}} + \frac{1}{2}M_2^4(\delta g^{00})^2$, can be treated perturbatively.

C.1 Preliminaries

Our starting point will be the low-energy effective action for the Goldstone mode in the decoupling limit

$$\mathcal{L}_{\text{eff}} = M_{\text{pl}}^2 |\dot{H}| \left[-(\partial_\mu \pi)^2 + \frac{\rho^2}{\mu^2} \left(\dot{\pi} - \frac{1}{2}(\partial_\mu \pi)^2 \right) \left(1 + \frac{\square}{\mu^2} + \dots \right) \left(\dot{\pi} - \frac{1}{2}(\partial_\mu \pi)^2 \right) \right] + \dots, \quad (\text{C.1})$$

where

$$c_s^{-2} \equiv 1 + \frac{\rho^2}{\mu^2}. \quad (\text{C.2})$$

We will compute correlation functions of the comoving curvature perturbation $\zeta = -H\pi$. The quadratic Lagrangian for ζ is

$$L_2 \equiv a^4 \mathcal{L}_2 = \mathcal{C} \cdot \frac{(aH)^2}{c_s^2} \left[(\zeta')^2 - c_s^2 (\partial_i \zeta)^2 \right], \quad (\text{C.3})$$

where $\mathcal{C} \equiv \frac{M_{\text{pl}}^2 |\dot{H}|}{H^4}$. For quasi-De Sitter backgrounds, $\mathcal{C} \approx \text{const.}$ and $(aH) \approx -\tau^{-1}$. The leading interactions in the limit $c_s \ll 1$ are

$$L_{\text{int}}^{(0)} = \mathcal{C} \cdot \frac{aH}{c_s^2} \cdot \zeta' (\partial_i \zeta)^2, \quad (\text{C.4})$$

$$L_{\text{int}}^{(1)} = \mathcal{C} \cdot \frac{H^2}{c_s^2 \mu^2} \cdot \zeta' \partial_i^2 \zeta', \quad (\text{C.5})$$

$$L_{\text{int}}^{(2)} = \mathcal{C} \cdot \frac{H^2}{c_s^2 \mu^2} \cdot \frac{aH}{c_s^2} \cdot \zeta' \frac{c_s^2 \partial_j^2}{(aH)^2} (\partial_i \zeta)^2. \quad (\text{C.6})$$

The interaction Hamiltonian is $H_{\text{int}} = -\int d^3x L_{\text{int}}$. We promote the field ζ to the operator $\hat{\zeta}$, whose Fourier modes we expand in creation and annihilation operators

$$\hat{\zeta}_{\mathbf{k}}(\tau) = \zeta_k(\tau) \hat{a}_{\mathbf{k}} + \zeta_k^*(\tau) \hat{a}_{-\mathbf{k}}^\dagger, \quad (\text{C.7})$$

where

$$[\hat{a}_{\mathbf{k}}, \hat{a}_{-\mathbf{k}'}^\dagger] = (2\pi)^3 \delta(\mathbf{k} + \mathbf{k}'). \quad (\text{C.8})$$

We implicitly treat $\hat{\zeta}$ as interaction picture fields whose time-evolution is determined by $H_0 = -\int d^3x L_2$. We will use the in-in formalism to compute correlation functions (for a recent review see [39])

$$\langle \hat{Q} \rangle(\tau) = \langle 0 | \left[\bar{T} e^{i \int_{-\infty}^{\tau} d\tau' \hat{H}_{\text{int}}(\tau')} \right] \hat{Q}(\tau) \left[T e^{-i \int_{-\infty}^{\tau} d\tau' \hat{H}_{\text{int}}(\tau')} \right] | 0 \rangle, \quad (\text{C.9})$$

where $|0\rangle$ is the vacuum of the free theory, $\hat{a}_{\mathbf{k}}|0\rangle \equiv 0$, and the symbols T and \bar{T} denote time-ordering and anti-time-ordering, respectively. To compute equal time n -point functions of ζ , we

let $\hat{Q} = \{ \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2}, \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}, \dots \}$. Expanding the exponentials in (C.9) in powers of \hat{H}_{int} allows us to compute correlation functions perturbatively. We evaluate each term in the series using contractions and normal ordering. After normal ordering, the only terms that are non-vanishing are those with all terms contracted. A contraction between two terms, $\hat{\zeta}_{\mathbf{k}}(\tau')$ (on the left) and $\hat{\zeta}_{\mathbf{q}}(\tau'')$ (on the right), gives

$$: \overline{\hat{\zeta}_{\mathbf{k}}(\tau') \hat{\zeta}_{\mathbf{q}}(\tau'')} : = [\hat{\zeta}_{\mathbf{k}}(\tau'), \hat{\zeta}_{\mathbf{q}}(\tau'')] = \zeta_k(\tau') \zeta_q^*(\tau'') \delta(\mathbf{k} + \mathbf{q}) . \quad (\text{C.10})$$

Feynman diagrams are a convenient way of keeping track of all necessary contractions. The final integrals will be highly oscillatory in the infinite past due to form of the mode functions $\propto e^{-ic_s k \tau}$. To evaluate the integral we therefore perform a Wick rotation $\tau \rightarrow i\tau$.

C.2 Two-Point Function

The free-field action (C.3) implies the standard mode functions in de Sitter space

$$\zeta_k(\tau) = \zeta_k^{(o)} \cdot e^{-ic_s k \tau} (1 + ic_s k \tau) , \quad (\text{C.11})$$

where

$$k^{3/2} \zeta_k^{(o)} \equiv \frac{H}{M_{\text{pl}}} \frac{i}{\sqrt{4c_s \epsilon}} . \quad (\text{C.12})$$

This allows us to compute the two-point function (or power spectrum) after horizon crossing, $P_\zeta(k) = |\zeta_k^{(o)}|^2$, where

$$\langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \rangle = (2\pi)^3 P_\zeta(k_1) \delta(\mathbf{k}_1 + \mathbf{k}_2) . \quad (\text{C.13})$$

The dimensionless power spectrum is

$$\Delta_\zeta \equiv k^3 P_\zeta = \frac{1}{4} \frac{H^2}{M_{\text{pl}}^2 c_s \epsilon} = \frac{1}{4} \frac{1}{\mathcal{C} c_s} . \quad (\text{C.14})$$

From the quadratic correction term $H_{\text{int}}^{(1)}$ we get a tree-level correction to the power spectrum

$$\lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \rangle(\tau) = -i \int_{-\infty}^0 d\tau' \langle [\hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2}(0), \hat{H}_{\text{int}}^{(1)}(\tau')] \rangle , \quad (\text{C.15})$$

or

$$\Delta P_\zeta(k) = -i \cdot \mathcal{C} \cdot \frac{H^2}{c_s^2 \mu^2} \times \zeta_k^{(o)} \zeta_k^{(o)} \times 2 \int_{-\infty}^0 d\tau' \left[k^2 \frac{d\zeta_k^*}{d\tau'} \frac{d\zeta_k^*}{d\tau'} \right] + c.c. \quad (\text{C.16})$$

Using the mode functions (C.11) and performing a Wick rotation to regulate the integral, we find

$$\Delta P_\zeta = \frac{1}{4} \frac{H^2}{\mu^2 c_s^2} \times P_\zeta . \quad (\text{C.17})$$

We deduce that $H_{\text{int}}^{(1)}$ simply induces an unobservable shift of the amplitude of the power spectrum.

C.3 Three-Point Function

Next, we compute the three-point function (or bispectrum)

$$\langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle \equiv (2\pi)^3 B_\zeta(k_1, k_2, k_3) \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) . \quad (\text{C.18})$$

The leading bispectrum, corresponding to the ‘pure c_s -theory’, is

$$\lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) = -i \int_{-\infty}^0 d\tau' \langle [\hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0), \hat{H}_{\text{int}}^{(0)}(\tau')] \rangle . \quad (\text{C.19})$$

Substituting $\hat{H}_{\text{int}}^{(0)}$, we find

$$\begin{aligned} \lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) &= i \cdot \mathcal{C} \cdot \frac{1}{c_s^2} \cdot (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \zeta_{k_1}^{(o)} \zeta_{k_2}^{(o)} \zeta_{k_3}^{(o)} \times \\ &\quad \frac{1}{2}(k_1^2 - k_2^2 - k_3^2) \times \int \frac{d\tau'}{\tau'} (\zeta_{k_1}^*)' \zeta_{k_2}^* \zeta_{k_3}^* + \text{perms.} + \text{c.c.} \end{aligned} \quad (\text{C.20})$$

This gives [23]

$$B_\zeta^{(0)} = \frac{1}{4} \frac{\Delta_\zeta^2}{c_s^2} \cdot \frac{K_1^6 - 3K_1^4 K_2^2 + 11K_1^3 K_3^3 - 4K_1^2 K_2^4 - 4K_1 K_2^2 K_3^3 + 12K_3^6}{K_3^9 K_1^3} , \quad (\text{C.21})$$

where

$$K_1 = k_1 + k_2 + k_3 , \quad (\text{C.22})$$

$$K_2 = (k_1 k_2 + k_2 k_3 + k_3 k_1)^{1/2} , \quad (\text{C.23})$$

$$K_3 = (k_1 k_2 k_3)^{1/3} . \quad (\text{C.24})$$

To compute the correction induced by a combination of $H_{\text{int}}^{(0)}$ and $H_{\text{int}}^{(1)}$, we find it convenient to use an alternative form of (C.9),

$$\begin{aligned} \lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) &= \int_{-\infty}^0 d\tau' \int_{-\infty}^{\tau'} d\tau'' \langle 0 | \hat{H}_{\text{int}}(\tau') \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0) \hat{H}_{\text{int}}(\tau'') | 0 \rangle \\ &\quad - \int_{-\infty}^0 d\tau' \int_{-\infty}^{\tau'} d\tau'' \langle 0 | \hat{H}_{\text{int}}(\tau'') \hat{H}_{\text{int}}(\tau') \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0) | 0 \rangle \\ &\quad - \int_{-\infty}^0 d\tau' \int_{-\infty}^{\tau'} d\tau'' \langle 0 | \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0) \hat{H}_{\text{int}}(\tau') \hat{H}_{\text{int}}(\tau'') | 0 \rangle + \dots , \end{aligned} \quad (\text{C.25})$$

where $\hat{H}_{\text{int}} \equiv \hat{H}_{\text{int}}^{(0)} + \hat{H}_{\text{int}}^{(1)}$. The leading corrections come from terms with one factor of $\hat{H}_{\text{int}}^{(1)}$ and one factor of $\hat{H}_{\text{int}}^{(2)}$. The first integral in (C.25) therefore is

$$\begin{aligned} \int_{-\infty}^0 d\tau' \int_{-\infty}^{\tau'} d\tau'' \langle 0 | \hat{H}_{\text{int}}^{(0)}(\tau') \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0) \hat{H}_{\text{int}}^{(1)}(\tau'') | 0 \rangle + \text{c.c.} &= \\ = \frac{\mathcal{C}^2}{c_s^2} \cdot \frac{H^2}{c_s^2 \mu^2} \cdot (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \zeta_{k_1}^{(o)} \zeta_{k_2}^{(o)} \zeta_{k_3}^{(o)} \times \\ \frac{1}{2}(k_3^2 - k_1^2 - k_2^2) k_3^2 \times \int_{-\infty}^0 \frac{d\tau'}{\tau'} \zeta_{k_1} \zeta_{k_2} \zeta_{k_3}' \int_{-\infty}^0 d\tau'' (\zeta_{k_3}')^* (\zeta_{k_3}')^* + \text{perms.} + \text{c.c.} \end{aligned} \quad (\text{C.26})$$

The second integral in (C.25) is

$$\begin{aligned}
& - \int_{-\infty}^0 d\tau' \int_{-\infty}^{\tau'} d\tau'' \langle 0 | \left[\hat{H}_{\text{int}}^{(0)}(\tau'') \hat{H}_{\text{int}}^{(1)}(\tau') + \hat{H}_{\text{int}}^{(1)}(\tau'') \hat{H}_{\text{int}}^{(0)}(\tau') \right] \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0) | 0 \rangle = \quad (\text{C.27}) \\
& = \frac{\mathcal{C}^2}{c_s^2} \cdot \frac{H^2}{c_s^2 \mu^2} \cdot (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \zeta_{k_1}^{(o)} \zeta_{k_2}^{(o)} \zeta_{k_3}^{(o)} \times \frac{1}{2} (k_3^2 - k_1^2 - k_2^2) k_3^2 \times \\
& \left[\int_{-\infty}^0 d\tau' (\zeta'_{k_3})^* \zeta'_{k_3} \int_{-\infty}^{\tau'} \frac{d\tau''}{\tau''} \zeta_{k_1} \zeta_{k_2} \zeta'_{k_3} + \int_{-\infty}^0 \frac{d\tau'}{\tau'} \zeta_{k_1} \zeta_{k_2} (\zeta'_{k_3})^* \int_{-\infty}^{\tau'} d\tau'' \zeta'_{k_3} \zeta'_{k_3} + \text{perms.} \right].
\end{aligned}$$

Finally, the third integral in (C.25) is the complex conjugate of the second integral. The sum of the three integrals has a rather complex analytic answer, $B_\zeta^{(1)}$. Showing this answer here wouldn't be very illuminating.

Finally, there is a tree-level correction to the bispectrum from the interaction $H_{\text{int}}^{(2)}$,

$$\lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) = -i \int_{-\infty}^0 d\tau' \langle [\hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0), \hat{H}_{\text{int}}^{(2)}(\tau')] \rangle. \quad (\text{C.28})$$

Substituting $\hat{H}_{\text{int}}^{(2)}$, we find

$$\begin{aligned}
\lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) &= i \cdot \mathcal{C} \cdot \frac{H^2}{c_s^2 \mu^2} \cdot (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \zeta_{k_1}^{(o)} \zeta_{k_2}^{(o)} \zeta_{k_3}^{(o)} \times \quad (\text{C.29}) \\
&\frac{1}{2} (k_1^2 - k_2^2 - k_3^2) k_1^2 \times \int_{-\infty}^0 d\tau' \tau' (\zeta_{k_1}^*)' \zeta_{k_2}^* \zeta_{k_3}^* + \text{perms.} + c.c.
\end{aligned}$$

Performing the same manipulations as before, we get

$$\begin{aligned}
B_\zeta^{(2)} &= \frac{1}{18} \frac{H^2}{c_s^2 \mu^2} \frac{\Delta_\zeta^2}{c_s^2} \cdot \frac{K_1^8 - 3K_1^6 K_2^2 - 7K_1^4 K_2^4 + 12K_1^2 K_2^6 + 17K_1^5 K_3^3}{K_3^9 K_1^5} \times \\
&\times \frac{-43K_1^3 K_2^2 K_3^3 + 36K_1 K_2^4 K_3^3 + 66K_1^2 K_3^6 - 48K_2^2 K_3^6}{K_3^9 K_1^5}. \quad (\text{C.30})
\end{aligned}$$

In §4.2 we discuss the shapes of all bispectra computed in this appendix.

D Small μ Limit of the π - σ Model

In this appendix, we compute the power spectrum and the bispectrum in the $\mu \rightarrow 0$ limit of the π - σ model. As we showed in Appendix B, the dynamics of the Goldstone mode in this limit are characterized by a non-linear dispersion relation $\omega \sim k^2/\rho$. The theory is therefore similar, but, as we will show, not identical to ghost inflation [22]. We also argued in Appendix B that the theory at the energies relevant for inflation, $\omega \sim H \ll \rho$, is described by a single degree of freedom. However, the single-field action in this limit is non-local, so we will find it more convenient to consider the local two-field action

$$\mathcal{L} \approx \rho \dot{\pi}_c \sigma - \frac{1}{2} \frac{(\partial_i \pi_c)^2}{a^2} - \frac{1}{2} \frac{(\partial_i \sigma)^2}{a^2} - \underbrace{\frac{1}{\xi} (\partial_\mu \pi_c)^2 \sigma}_{\mathcal{L}_{\text{int}}} , \quad (\text{D.1})$$

where $\xi \equiv \frac{2M_{\text{pl}}^2 |\dot{H}|}{m^3} = \frac{2}{\rho} (2M_{\text{pl}}^2 |\dot{H}|)^{1/2}$. The quadratic part of the action, \mathcal{L}_2 , implies the following equations of motion

$$\ddot{\pi}_c + 5H\dot{\pi}_c + \frac{k^4}{\rho^2 a^4} \pi_c = 0 \quad \text{and} \quad \frac{k^2}{a^2} \sigma = \rho \dot{\pi}_c . \quad (\text{D.2})$$

In this appendix, we quantize the theory and compute the power spectrum and bispectrum of curvature fluctuations $\zeta = -H\pi$.

D.1 Canonical Quantization

At low energies, $\omega \ll \rho$, the coupling between π_c and σ dominates the dynamics. The field σ then plays the role of the conjugate momentum of π_c ,

$$p_\pi \equiv \frac{\partial \mathcal{L}'}{\partial \dot{\pi}_c} = \rho \sigma . \quad (\text{D.3})$$

The equal-time canonical commutation relation

$$[\hat{\pi}_c(\mathbf{x}, \tau), \hat{p}_\pi(\mathbf{y}, \tau)] = i a^{-3} \delta(\mathbf{x} - \mathbf{y}) , \quad (\text{D.4})$$

then implies

$$[\hat{\pi}_c(\mathbf{k}, \tau), \hat{\sigma}(\mathbf{q}, \tau)] = \frac{i}{a^3 \rho} (2\pi)^3 \delta(\mathbf{k} + \mathbf{q}) . \quad (\text{D.5})$$

This confirms that the fields π_c and σ are not independent degrees of freedom at low energies. In particular, at low energies, σ is proportional to π'_c ,

$$\frac{k^2}{a^2} \sigma = \frac{\rho}{a} \pi'_c . \quad (\text{D.6})$$

To obtain an analytic solution to the equation of motion (D.2), we use conformal time and define $v \equiv a^2 \pi_c$, such that

$$v'' + \left(\kappa^4 \tau^2 - \frac{6}{\tau^2} \right) v = 0 , \quad \text{where} \quad \kappa^2 \equiv \frac{k^2 H}{\rho} . \quad (\text{D.7})$$

This has a solution in terms of Hankel functions,

$$v = (-\tau)^{1/2} \left[c_1(\kappa) H_{5/4}^{(1)} \left(\frac{1}{2}(\kappa\tau)^2 \right) + c_2(k) H_{5/4}^{(2)} \left(\frac{1}{2}(\kappa\tau)^2 \right) \right] . \quad (\text{D.8})$$

Our unusual commutation relations imply that some care is required to define the correct Bunch-Davies vacuum. From (D.5) and (D.6) we obtain

$$[\hat{v}(\mathbf{k}, \tau), \hat{v}'(\mathbf{q}, \tau)] = i \frac{k^2}{\rho^2} (2\pi)^3 \delta(\mathbf{k} + \mathbf{q}) . \quad (\text{D.9})$$

This implies a non-standard normalization for the mode functions

$$v_k v_k^{*'} - v_k^* v_k' = i \frac{k^2}{\rho^2} , \quad (\text{D.10})$$

if we define the standard operator mode expansion

$$\hat{v}(\mathbf{k}, \tau) = v_k(\tau) \hat{a}_{\mathbf{k}} + v_k^*(\tau) \hat{a}_{-\mathbf{k}}^\dagger , \quad (\text{D.11})$$

with $[a_{\mathbf{k}}, a_{-\mathbf{q}}^\dagger] = (2\pi)^3 \delta(\mathbf{k} + \mathbf{q})$. Substituting (D.8) into (D.10) gives

$$\frac{8}{\pi} \left[|c_1(k)|^2 - |c_2(k)|^2 \right] = \frac{k^2}{\rho^2} . \quad (\text{D.12})$$

Selecting the positive frequency solutions at early times fixes $c_2 = 0$, and hence

$$c_1(k) = \sqrt{\frac{\pi}{8}} \frac{k}{\rho} . \quad (\text{D.13})$$

The mode functions are then completely determined

$$v_k(\tau) = \sqrt{\frac{\pi}{8}} \frac{k}{\rho} (-\tau)^{1/2} H_{5/4}^{(1)} \left(\frac{1}{2} \frac{H}{\rho} (k\tau)^2 \right) . \quad (\text{D.14})$$

In the superhorizon limit this becomes

$$v_k^{(o)} \equiv \lim_{k\tau \rightarrow 0} v_k \sim a^2 k^{-3/2} (\rho H^3)^{1/4} . \quad (\text{D.15})$$

D.2 Two-Point Function

Using $v^2 = a^4 \pi_c^2 = 2a^4 M_{\text{pl}}^2 \epsilon \zeta^2$, we arrive at the power spectrum for ζ after horizon crossing

$$\Delta_\zeta = k^3 |\zeta_k^{(o)}|^2 \sim \frac{H^2}{M_{\text{pl}}^2 \epsilon} \left(\frac{\rho}{H} \right)^{1/2} . \quad (\text{D.16})$$

We note that the spectrum is scale-invariant. The amplitude is enhanced by a factor of $\rho/H \gg 1$ relative to the familiar expression from slow-roll inflation.

D.3 Three-Point Function

The three-point function in the small μ limit is quite interesting. The interactions involve new operators that contain both π and σ fields. Because σ is the canonical momentum, it is not clear what to expect from these interactions.

A rough measure of the expected size of non-Gaussianity is

$$\frac{\mathcal{L}_3}{\mathcal{L}_2} \sim \frac{\xi^{-1}(\partial_\mu \pi_c)^2 \sigma}{\rho \dot{\pi}_c \sigma} \sim \frac{1}{\xi \rho} \frac{(\partial_i \pi_c)^2}{\dot{\pi}_c} \sim \frac{k^2}{\omega} \frac{\pi_c}{(M_{\text{pl}}^2 |\dot{H}|)^{1/2}} \sim \rho \pi \sim \frac{\rho}{H} \zeta, \quad (\text{D.17})$$

or

$$f_{\text{NL}} \sim \frac{1}{\zeta} \frac{\mathcal{L}_3}{\mathcal{L}_2} \sim \frac{\rho}{H} \gg 1. \quad (\text{D.18})$$

In the regime of interest this is a detectably large amount of non-Gaussianity. We compute the detailed shape of the bispectrum using the formalism of Appendix C,

$$\lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle = -i \int_{-\infty}^0 d\tau' \langle [\hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3}(0), \hat{H}_{\text{int}}(\tau')] \rangle, \quad (\text{D.19})$$

where

$$H_{\text{int}} = - \int d^3x a^4 \mathcal{L}_{\text{int}} \approx - \frac{2\mathcal{C}}{\xi} \int d^3x (aH)^2 (\partial_i \zeta)^2 \sigma. \quad (\text{D.20})$$

The bispectrum is then determined by the following integral

$$\begin{aligned} \lim_{\tau \rightarrow 0} \langle \hat{\zeta}_{\mathbf{k}_1} \hat{\zeta}_{\mathbf{k}_2} \hat{\zeta}_{\mathbf{k}_3} \rangle(\tau) &= i \cdot \mathcal{C} \cdot \frac{\rho^2}{H^2} \cdot (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3) \times \zeta_{k_1}^{(o)} \zeta_{k_2}^{(o)} \zeta_{k_3}^{(o)} \times \\ &\quad \frac{\frac{1}{2}(k_1^2 - k_2^2 - k_3^2)}{k_1^2} \times \int_{-\infty}^0 \frac{d\tau'}{(\tau')^3} (\zeta_{k_1}^*)' \zeta_{k_2}^* \zeta_{k_3}^* + \text{perms.} + c.c. \end{aligned} \quad (\text{D.21})$$

where we used

$$\frac{\sigma_k}{\xi} = \frac{1}{2} \frac{aH}{k^2} \cdot \frac{\rho^2}{H^2} \cdot \zeta_k'. \quad (\text{D.22})$$

We find that the shape generated by this interaction is almost indistinguishable from the equilateral shape generated by $M_2^4 \dot{\pi} (\partial_i \pi)^2$ in the small c_s -theory (see §4.4).

E Unitarity Bounds

In this appendix, we derive unitarity bounds for the models discussed in the paper.

E.1 Preliminaries

Unitarity of the S-matrix translates into a statement about amplitudes through the optical theorem [40]:

$$\begin{aligned} 2 \operatorname{Im} \mathcal{A}(k_1, k_2 \rightarrow p_1, p_2) &= \\ &= \prod_i \int \frac{d^3 q_i}{(2\pi)^3} \frac{1}{2\omega_i} \mathcal{A}(k_1, k_2 \rightarrow \{q_i\}) \mathcal{A}^*(p_1, p_2 \rightarrow \{q_i\}) \delta^{(4)}\left(k_1 + k_2 - \sum_i q_i\right). \end{aligned} \quad (\text{E.1})$$

This expression will be most useful for $2 \rightarrow 2$ scattering, in which case the matrix elements on both sides of the optical theorem are the same. It is convenient to rewrite the amplitude for $2 \rightarrow 2$ scattering in the center of mass frame and expand it in Legendre polynomials

$$\mathcal{A}(k_1, k_2 \rightarrow p_1, p_2)_{\text{cm}} = \left(\frac{\partial k}{\partial \omega} \frac{k^2}{\omega^2} \right)^{-1} 16\pi \sum_{\ell} (2\ell + 1) P_{\ell}(\cos \theta) a_{\ell}, \quad (\text{E.2})$$

where $\cos \theta = \frac{1}{2} \hat{\mathbf{p}}_1 \cdot \hat{\mathbf{p}}_2$. The subscript ℓ labels the spin- ℓ partial wave. The prefactor in brackets is unity in a relativistic theory, but will play an important role in our non-relativistic examples. Due to angular momentum conservation, we can write (E.1) as

$$\operatorname{Im} a_{\ell} = |a_{\ell}|^2. \quad (\text{E.3})$$

This expression leads to a powerful constraint, because it cannot be satisfied if $(\operatorname{Re} a_{\ell})^2 > \frac{1}{4}$. A violation of (E.3) implies a violation of unitarity. In practice, the leading contributions to $\operatorname{Re} a_{\ell}$ are computed perturbatively. When these contributions to $\operatorname{Re} a_{\ell}$ exceed $\frac{1}{2}$, one concludes that the theory is strongly coupled, i.e. higher-order terms must be of equal importance for the result to be consistent with unitarity. Let's see how this works in our examples.

E.2 Unitarity and Small Sound Speed

We will start with the ‘pure c_s -theory’, $\mathcal{L}_{\text{sr}} + \frac{1}{2} M_2^4 (\delta g^{00})^2$. We will consider the unitarity bound associated with the operator $\frac{1}{2} M_2^4 (\partial_i \pi \partial^i \pi)^2$ which, after canonical-normalization, becomes

$$\mathcal{L}_{\text{int}} = \frac{1}{8} \frac{(1 - c_s^2) c_s^2}{2M_{\text{pl}}^2 |\dot{H}|} (\partial_i \pi_c \partial^i \pi_c)^2. \quad (\text{E.4})$$

For convenience, we have extracted a factor of $\frac{1}{8}$ since it equals the combinatorial factor arising from this interaction. The strength of the interaction is controlled by the dimensionless ratio $\omega^4 / (2M_{\text{pl}}^2 \dot{H} c_s^5)$. We expect the theory to become strongly coupled at some order-one value of this ratio. We will use the perturbative violation of (E.3) to compute this number.

The amplitude generated by the interaction (E.4) is

$$\mathcal{A}(k_1, k_2 \rightarrow p_1, p_2)_{\text{cm}} = \frac{(1 - c_s^2)}{2M_{\text{pl}}^2 |\dot{H}| c_s^2} \omega^4 [1 + 2 \cos^2(\theta)]. \quad (\text{E.5})$$

We compare this to (E.2), which now reads

$$\mathcal{A}(k_1, k_2 \rightarrow p_1, p_2)_{\text{cm}} = c_s^3 \cdot 16\pi \sum_{\ell} (2\ell + 1) P_{\ell}(\cos \theta) a_{\ell} . \quad (\text{E.6})$$

The largest partial wave amplitude is the s-wave component

$$a_0 = \frac{1}{16\pi} \frac{(1 - c_s^2)\omega^4}{2M_{\text{pl}}^2 |\dot{H}| c_s^5} \int d\cos\theta [1 + 2\cos^2\theta] = \frac{1}{4\pi} \frac{(1 - c_s^2)\omega^4}{2M_{\text{pl}}^2 |\dot{H}| c_s^5} . \quad (\text{E.7})$$

Using $a_0 < \frac{1}{2}$ we find

$$\omega^4 < 4\pi M_{\text{pl}}^2 |\dot{H}| \frac{c_s^5}{1 - c_s^2} . \quad (\text{E.8})$$

E.3 The Extrinsic Curvature Model

Let us apply the same reasoning to our UV-completion with extrinsic curvature terms (cf. §3.2.2). As in the small c_s -theory, the strong coupling scale is determined by the operator $\frac{1}{2}M_2^4(\partial_i\pi\partial^i\pi)^2$. After canonical-normalization, the interaction is

$$\mathcal{L}_{\text{int}} = \frac{1}{8} \frac{1}{4M_2^4} (\partial_i\pi_c\partial^i\pi_c)^2 . \quad (\text{E.9})$$

Due to the modified dispersion relation, the strength of the interaction is controlled by the dimensionless ratio $\omega^{1/2}/(M_2^4\rho^{7/2})$. We will compute the order-one value of this ratio at the energy scale at which the theory becomes strongly coupled.

The amplitude for $2 \rightarrow 2$ scattering generated by the interaction (E.9) is

$$\mathcal{A}(k_1, k_2 \rightarrow p_1, p_2)_{\text{cm}} = \frac{\rho^2}{4M_2^4} \omega^2 [1 + 2\cos^2\theta] . \quad (\text{E.10})$$

This answer is essentially the same as in the small c_s -case, except that we have used a different dispersion relation to write the amplitude in terms of ω alone. Again, the largest partial wave amplitude is the s-wave component

$$a_0 = \frac{1}{16\pi} \frac{\rho^{7/2}\omega^{1/2}}{M_2^4} . \quad (\text{E.11})$$

Using (E.3), we require $a_0 < 1/2$ for the S-matrix to be unitary. This implies

$$\omega < (8\pi)^2 \frac{M_2^8}{\rho^7} . \quad (\text{E.12})$$

The numerical factor in (E.12) is larger than in the ‘pure c_s -theory’ because the operator has a smaller scaling dimension.

E.4 The π - σ Model

Finally, we wish to determine the scale at which the π - σ model (§3.2.1) becomes strongly coupled. Unlike the previous two examples, perturbative unitarity is not a useful measure of strong coupling. As we will show, infrared divergences associated with the massless Goldstone bosons make the analysis ill-defined. We will instead define strong coupling by estimating the size of loop corrections directly.

The largest coupling in the π - σ model arises from the interaction $m^3 \partial_i \pi \partial^i \pi \sigma$. After canonical-normalization, the interaction becomes

$$\mathcal{L}_{\text{int}} = \frac{1}{2} \frac{1}{(2M_{\text{pl}}|\dot{H}|)^{1/2} \rho^{1/2}} \partial_i \pi_c \partial^i \pi_c \sigma_c . \quad (\text{E.13})$$

We are interested in $2 \rightarrow 2$ scattering of the Goldstone bosons. Because σ is not an independent field, we need to be careful about properly defining the propagator. Completing the square as usual, we find that the propagator takes the form

$$\langle \Psi^i \Psi_j \rangle = \frac{1}{\omega^2 - k^4/\rho^2} \begin{pmatrix} \frac{k^2}{\rho} & -i\omega \\ i\omega & \frac{k^2}{\rho} \end{pmatrix} , \quad (\text{E.14})$$

where $\Psi^i \equiv (\sigma \pi)$. We note that on shell $\sigma = \rho \dot{\pi}/k^2$. Since the external lines are on shell, the Feynman rules for the external lines involving σ are identical to those of π . Using Feynman rules, we can therefore compute the $2 \rightarrow 2$ scattering amplitude. The full amplitude is not important, as we only need to understand the $\theta \rightarrow 0$ limit

$$\mathcal{A}(k_1, k_2 \rightarrow p_1 = k_1, p_2 = k_2)_{\text{cm}} \simeq \frac{2}{\theta^2} \frac{\omega^3 \rho}{M_{\text{pl}}^2 |\dot{H}|} . \quad (\text{E.15})$$

The divergence as $\theta \rightarrow 0$ is associated with the t-channel exchange of a massless particle with vanishing energy and momentum. This divergence is common in a theory with massless particles and reflects the fact that the probability for final states with a finite number of particles is zero. Unfortunately, the divergence also undermines our ability to use the partial wave expansion to determine the strong coupling scale. These infrared divergences infect our observables and prevent us from isolating the UV-behavior.

We will therefore define strong coupling as the scale where one-loop and tree-level contributions to the same process are equal. Because (E.13) is a three-particle interaction, for every loop integral we need two insertions of the coupling, and hence get a suppression by $16\pi^2$. Finally, we get the following bound

$$\omega^{3/2} < 16\pi^2 \frac{2M_{\text{pl}}^2 |\dot{H}|}{\rho^{5/2}} . \quad (\text{E.16})$$

The additional powers of ρ come from the momentum integrals in every loop. This is equivalent to defining a dimensionless coupling by using the dispersion relation.

References

- [1] S. Weinberg, “The Quantum Theory of Fields. Vol. 2: Modern Applications,” *Cambridge, UK: Univ. Pr. (1996)* 489 p
- [2] S. Weinberg, “Phenomenological Lagrangians,” *Physica* **A96**, 327 (1979).
- [3] C. Cheung *et al.*, “The Effective Field Theory of Inflation,” *JHEP* **0803**, 014 (2008).
- [4] P. Creminelli *et al.*, “Starting the Universe: Stable Violation of the Null Energy Condition and Non-Standard Cosmologies,” *JHEP* **0612**, 080 (2006).
- [5] L. Senatore and M. Zaldarriaga, “The Effective Field Theory of Multi-Field Inflation,” arXiv:1009.2093 [hep-th].
- [6] B. Lee, C. Quigg and H. Thacker, “Weak Interactions At Very High-Energies: The Role Of The Higgs Boson Mass,” *Phys. Rev. D* **16**, 1519 (1977);
B. Lee, C. Quigg and H. Thacker, “The Strength Of Weak Interactions At Very High-Energies And The Higgs Boson Mass,” *Phys. Rev. Lett.* **38**, 883 (1977).
- [7] A. Guth, “The Inflationary Universe: A Possible Solution To The Horizon And Flatness Problems,” *Phys. Rev. D* **23**, 347 (1981);
A. Linde, “A New Inflationary Universe Scenario: A Possible Solution Of The Horizon, Flatness, Homogeneity, Isotropy And Primordial Monopole Problems,” *Phys. Lett. B* **108**, 389 (1982);
A. Albrecht and P. Steinhardt, “Cosmology For Grand Unified Theories With Radiatively Induced Symmetry Breaking,” *Phys. Rev. Lett.* **48**, 1220 (1982);
for a recent review see e.g. D. Baumann, “TASI Lectures on Inflation,” arXiv:0907.5424 [hep-th].
- [8] E. Komatsu *et al.*, “Non-Gaussianity as a Probe of the Physics of the Primordial Universe and the Astrophysics of the Low Redshift Universe,” [arXiv:0902.4759 [astro-ph.CO]].
- [9] M. Peskin and T. Takeuchi, “Estimation of Oblique Electroweak Corrections,” *Phys. Rev. D* **46**, 381-409 (1992);
R. Barbieri *et al.*, “Electroweak Symmetry Breaking after LEP-1 and LEP-2,” *Nucl. Phys. B* **703**, 127-146 (2004).
- [10] P. Creminelli, “On Non-Gaussianities in Single-Field Inflation,” *JCAP* **0310**, 003 (2003).
- [11] S. Weinberg, “Effective Field Theory for Inflation,” *Phys. Rev. D* **77**, 123541 (2008).
- [12] E. Silverstein and D. Tong, “Scalar Speed Limits and Cosmology: Acceleration from D-cceleration,” *Phys. Rev. D* **70**, 103505 (2004).
- [13] C. Burrage *et al.*, “Galileon Inflation,” *JCAP* **1101**, 014 (2011).

- [14] X. Chen, “Inflation from Warped Space,” JHEP **0508**, 045 (2005);
R. Bean, X. Chen, H. Peiris, J. Xu, “Comparing Infrared Dirac-Born-Infeld Brane Inflation to Observations,” Phys. Rev. **D77**, 023527 (2008).
- [15] A. Tolley and M. Wyman, “The Gelaton Scenario: Equilateral Non-Gaussianity from Multi-Field Dynamics,” Phys. Rev. **D81**, 043502 (2010).
- [16] S. Cremonini, Z. Lalak, and K. Turzynski, “Strongly Coupled Perturbations in Two-Field Inflationary Models,” [arXiv:1010.3021 [hep-th]].
- [17] A. Achucarro *et al.*, “Features of Heavy Physics in the CMB Power Spectrum,” [arXiv:1010.3693 [hep-ph]].
- [18] X. Chen and Y. Wang, “Quasi-Single-Field Inflation and Non-Gaussianities,” JCAP **1004**, 027 (2010).
- [19] A. Berera, “Warm Inflation,” Phys. Rev. Lett. **75**, 3218 (1995).
- [20] D. Green *et al.*, “Trapped Inflation,” Phys. Rev. D **80**, 063533 (2009).
- [21] D. Nacir *et al.*, “Dissipative Effects in the Effective Field Theory of Inflation,” to appear.
- [22] N. Arkani-Hamed *et al.*, “Ghost Inflation,” JCAP **0404**, 001 (2004).
- [23] L. Senatore, K. Smith, and M. Zaldarriaga, “Non-Gaussianities in Single-Field Inflation and their Optimal Limits from the WMAP 5-year Data,” JCAP **1001**, 028 (2010).
- [24] J. Maldacena, “Non-Gaussian Features of Primordial Fluctuations in Single-Field Inflationary Models,” JHEP **0305**, 013 (2003).
- [25] S. Weinberg, “Adiabatic Modes in Cosmology,” Phys. Rev. **D67**, 123504 (2003).
- [26] R. Arnowitt, S. Deser, and C. Misner, “The Dynamics of General Relativity,” [gr-qc/0405109].
- [27] C. Cheung *et al.*, “On the Consistency Relation of the 3-Point Function in Single-Field Inflation,” JCAP **0802**, 021 (2008).
- [28] E. Komatsu *et al.*, “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation,” [arXiv:1001.4538 [astro-ph.CO]].
- [29] D. Babich, P. Creminelli, and M. Zaldarriaga, “The Shape of Non-Gaussianities,” JCAP **0408**, 009 (2004).
- [30] J. Fergusson and P. Shellard, “The Shape of Primordial Non-Gaussianity and the CMB Bispectrum,” Phys. Rev. **D80**, 043510 (2009).
- [31] N. Bartolo *et al.*, “Large non-Gaussianities in the Effective Field Theory Approach to Single-Field Inflation: the Bispectrum,” JCAP **1008**, 008 (2010).

- [32] M. Liguori *et al.*, “Primordial non-Gaussianity and Bispectrum Measurements in the Cosmic Microwave Background and Large-Scale Structure,” *Adv. Astron.* **2010**, 980523 (2010).
- [33] Amit Yadav, private communication.
- [34] The Planck Collaboration, “Planck Early Results: The Planck Mission,” [arXiv:1101.2022 [astro-ph.IM]].
- [35] D. Baumann *et al.*, “CMBPol Mission Concept Study: Probing Inflation with CMB Polarization,” *AIP Conf. Proc.* **1141**, 10 (2009).
- [36] The CORe Collaboration, “CORe (Cosmic Origins Explorer) A White Paper,” [arXiv:1102.2181 [astro-ph.CO]].
- [37] L. Senatore and M. Zaldarriaga, “A Naturally Large Four-Point Function in Single-Field Inflation,” [arXiv:1004.1201 [hep-th]].
- [38] A. Tseytlin, “Born-Infeld Action, Supersymmetry and String Theory,” arXiv:hep-th/9908105.
- [39] X. Chen, “Primordial Non-Gaussianities from Inflation Models,” *Adv. Astron.* **2010**, 638979 (2010).
- [40] M. Peskin and D. Schroeder, “An Introduction to Quantum Field Theory,” Reading, USA: Addison-Wesley (1995) 842 p.